



SCIENTIFIC
AMERICAN

12 Revolutionary Discoveries That Could Change Everything

Course Guidebook

Laura Helmuth



Copyright © The Teaching Company, 2024

Printed in the United States of America

This book is in copyright. All rights reserved.

Without limiting the rights under copyright reserved above, no part of this publication may be reproduced, stored in or introduced into a retrieval system, or transmitted, in any form, or by any means (electronic, mechanical, photocopying, recording, or otherwise), without the prior written permission of The Teaching Company.

4840 Westfields Boulevard, Suite 400

Chantilly, VA 20151-2299

USA

1-800-832-2412

www.thegreatcourses.com



Laura Helmuth is the editor in chief of *Scientific American*. She earned a PhD in Cognitive Neuroscience from the University of California, Berkeley, and a graduate certificate in Science Communication from the University of California, Santa Cruz. She has served in various editorial positions, including for *The Washington Post*, *Smithsonian* magazine, and *Science*. She is a member of the National Academies of Sciences, Engineering, and Medicine’s Standing Committee on Advancing Science Communication and was previously president of the National Association of Science Writers.

SCIENTIFIC AMERICAN

About Our Partner

Scientific American covers the advances in research and discovery that are changing our understanding of the world and shaping our lives. Founded 1845, it is the oldest continuously published magazine in the United States and now reaches more than 10 million people around the world each month through its website, print and digital editions, newsletters, and app. Authoritative and engaging features, news, opinion, and multimedia stories from journalists and expert authors—including more than 200 Nobel Prize winners—provide need-to-know coverage, insights, and illumination of the most important developments at the intersection of science and society. *Scientific American* is published by Springer Nature. As a research publisher, Springer Nature is home to other trusted brands, including Springer, Nature Research, BMC, and Palgrave Macmillan.

Based on several articles from Revolutionary Science—a collector's edition of *Scientific American*—this course highlights areas where new research and discoveries have reshaped scientists' understanding of the world and raised new questions. For more information on specific topics within this course, refer to the course scope on page 1.

Table of Contents

About Laura Helmuth	i
About Our Partner	iii
Course Scope	1
1. Why the Sea Can Glow the Color of Milk	2
2. The True Colors of Dinosaurs	10
3. How Textiles Gave Viking Women Power	20
4. New Truths about Lichens	32
5. The Quantum Nature of Bird Migration	42
6. Unexpected Smashups in the Solar System	54
7. Why Exercise Doesn't Help You Lose Weight	64
8. How Parents' Trauma Affects Future Children	74
9. The Breakthrough in Messenger RNA Therapies	86
10. You Really Can Learn in Your Sleep	90
11. How Your Brain Constructs the World	98
12. Why Black Holes Turn Inside Out	108

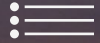
Course Scope

This course's content was adapted from a collection of articles, many of which appear in the collector's edition of *Scientific American* titled Revolutionary Science. The articles share a theme of surprising discoveries or research findings that have challenged previously held notions in a particular field. The course begins in the ocean, with the mysterious glow of "milky seas" that once puzzled mariners and has since been attributed to bioluminescent organisms. Lecture 2 looks at how fossilized pigments have revealed that many dinosaurs were colorful creatures. The third lecture turns to ancient cloth woven by Viking women and the story it tells about their power as weavers.

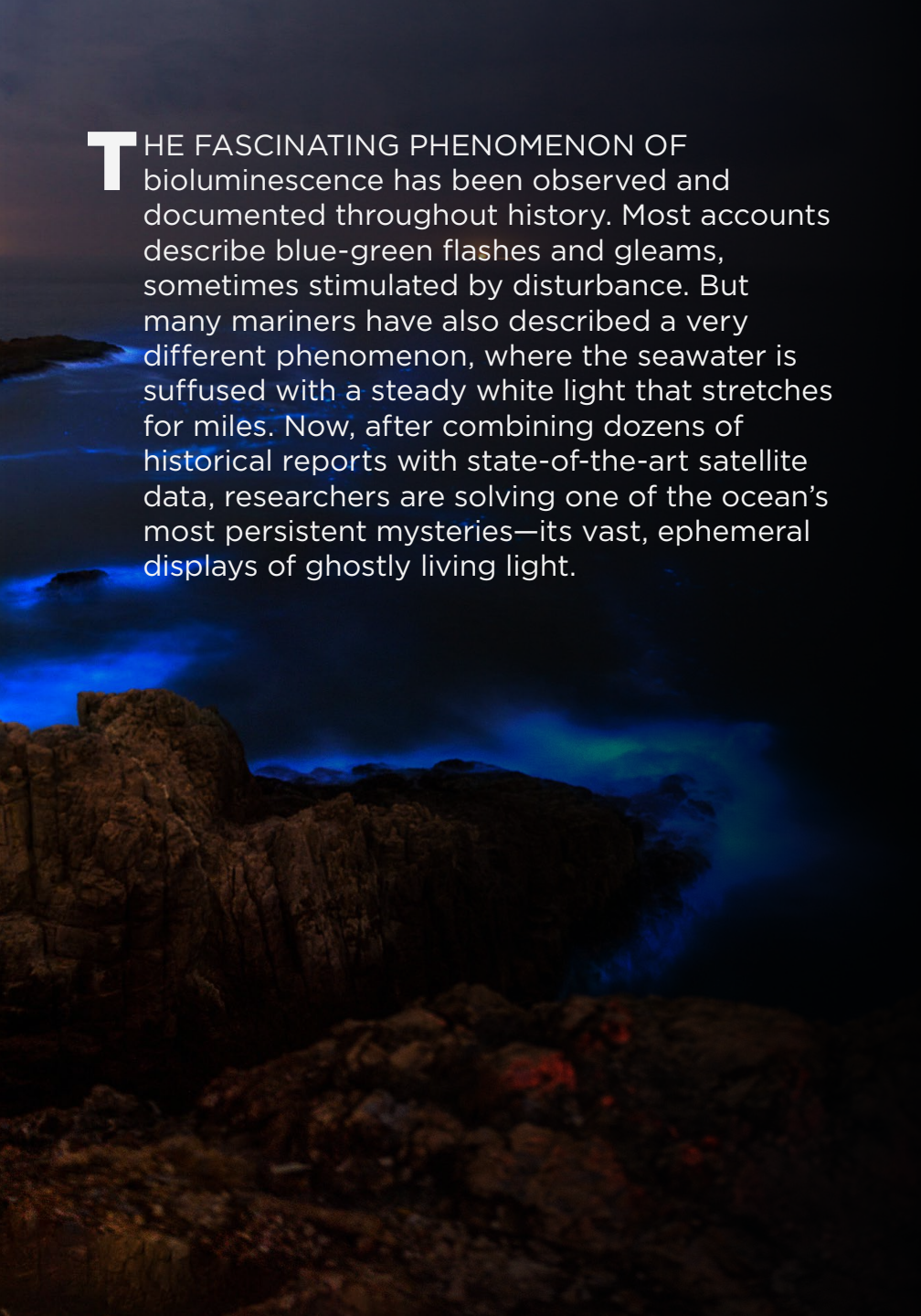
In lecture 4, researchers discuss how diverse perspectives led to richer findings in the study of lichens, and in lecture 5, scientists explain how birds use magnetoreception to navigate long-distance migrations. Lecture 6 shares the unexpected findings about the early solar system that have some scientists rethinking planetary evolution. Lecture 7 identifies unique features of human metabolism that explain why exercise is not the best approach to weight loss.

The next three lectures cover the following topics: the intergenerational effects of trauma (lecture 8), the promising results of mRNA therapies (lecture 9), and the critical relationship between sleep, memory, and learning (lecture 10). Lecture 11 reveals new findings in neuroscience that challenge the idea that the human mind begins as a blank slate. The course concludes in space, with new insights into the nature of black holes.

1



Why the Sea Can Glow the Color of Milk

A photograph of a rocky coastline at night. The foreground shows dark, jagged rocks. In the middle ground, the ocean is illuminated with a vibrant blue and green glow, likely from bioluminescence. The background is a dark, starry sky.

THE FASCINATING PHENOMENON OF bioluminescence has been observed and documented throughout history. Most accounts describe blue-green flashes and gleams, sometimes stimulated by disturbance. But many mariners have also described a very different phenomenon, where the seawater is suffused with a steady white light that stretches for miles. Now, after combining dozens of historical reports with state-of-the-art satellite data, researchers are solving one of the ocean's most persistent mysteries—its vast, ephemeral displays of ghostly living light.

From Science Fiction to Science

The cold radiance emitted by fireflies, some species of fungi, and various sea creatures is called bioluminescence. In the late 1800s, after centuries of speculation, scientists confirmed that bioluminescence results from an oxidation reaction between an enzyme and its substrate within animal and plant cells. But basic questions remained, however: No one knew what prompted different organisms to glow or what purpose the light might serve.

These “milky seas” were rare enough, and strange enough, that people widely considered them to be tall tales—barely more plausible than mermaid encounters. In Jules Verne’s novel *Twenty Thousand Leagues under the Sea*, the fictional marine biologist Pierre Aronnax is less perturbed by his voyage through a milky sea in the Bay of Bengal. He calmly informs his assistant that “the whiteness which surprises you is caused only by the presence of myriads of infusoria, a sort of luminous little worm, gelatinous and without color.”

Verne’s pilot was on the right track, but it would be more than a century before science began to catch up with science fiction. In July 1985, a US Navy research vessel encountered a milky sea off the Arabian Peninsula. The scientists onboard, who were conducting a broad study of marine bioluminescence, were equipped for this stroke of luck, and they quickly collected seawater samples for inspection.

In addition to the dinoflagellates, copepods, and other types of plankton associated with the familiar flashing displays, the samples contained bioluminescent bacteria. The researchers suggested that milky seas occurred after algae colonies on the water’s surface bloomed and died. When the dead algal cells ruptured, they released lipids that were then consumed by bacteria. These bacteria multiplied furiously, eventually becoming concentrated enough to produce a continuous glow.

Finally, milky seas had been established as a scientific phenomenon with a biological cause. But to understand where, when, and exactly why they occurred, researchers needed more data than serendipity could provide.

Satellite Images of Milky Seas

For the US Navy, marine bioluminescence is a practical concern because a patch of bright seawater can outline a submarine, turning it into an easy target. In the early 2000s, Steven Miller, an atmospheric scientist then at the Naval Research Laboratory in Monterey, California, began to wonder whether satellite sensors could detect bioluminescence from above. The only sensors capable of observing visible light at night were those in the Operational Linescan System (or OLS) that flew on US Air Force satellites.

Miller knew that most surface displays of marine bioluminescence were much too small to register on the sensors, so, on a whim, he searched the internet for mentions of widespread bioluminescence. He turned up a description of milky seas on the website Science Frontiers, an idiosyncratic catalog of “unusual & unexplained” happenings.

Miller began to collect eyewitness accounts. Among them was a report from a British merchant ship, the SS *Lima*, which had sailed through a milky sea along the Horn of Africa on January 25, 1995. When Miller pulled up the OLS images from the *Lima*'s location on that date, he saw a faint, comma-shaped smear that “looked like a finger smudge.”

He found that the edges of the smudge matched the coordinates noted in the ship's log as it entered and exited the milky sea, which covered nearly 5,500 square miles. When he examined OLS images from the days immediately before and after the *Lima*'s encounter, he found the same smudge, rotating counterclockwise in concert with local ocean currents. “Okay,” Miller thought, “we can see bioluminescence from space.”

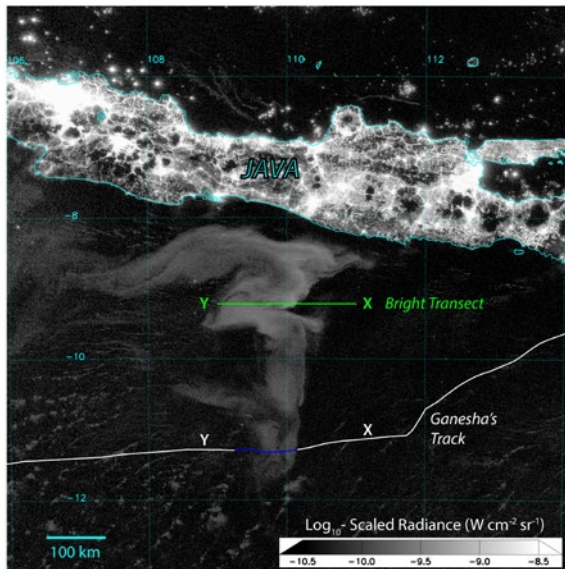
Miller got in touch with Steven Haddock, a marine biologist at the nearby Monterey Bay Aquarium Research Institute (known as MBARI), to share his findings. Haddock, who primarily studies bioluminescence in jellyfish, had spent much of his career trying to get as close as possible to bioluminescent organisms using crewed or remotely operated deep-sea submersibles. He and Miller began to collaborate.

The OLS detection from 1995 had been something of a fluke—the product of Miller’s persistence and a fortuitous satellite position. Miller hoped a new, more sensitive low-light visible-spectrum instrument called the day-night band (DNB) sensor would allow a systematic survey of milky seas. The sensor, launched in 2011, now rides on two satellites more than 500 miles above Earth’s surface, each orbiting the planet daily. More than 100 times as sensitive as the OLS, the DNB sensor can easily pick up the gleam of a milky sea.

Thanks to the long history of sightings by mariners, Miller and Haddock knew that the occurrences of milky seas peaked in winter and summer. They were most often reported in the northwestern Indian Ocean, as well as around Indonesia, particularly near the island of Java and in the Banda Sea.

Narrowing his search to these seasons and locations, Miller analyzed DNB data collected on moonless nights from 2012 through 2021. One 2019 event, detected just south of Java, was visible for at least 45 nights and covered almost 40,000 square miles—an area the size of Kentucky.

Miller confirmed the satellite observation with the first known photos of a milky seas event, taken by crew members of a ketch sailing near Java. The images confirmed the accuracy of Miller’s satellite detection methods and provided a long-sought glimpse of this elusive phenomenon.



Hypotheses for How and Why Milky Seas Form

From a distance, scientists have proposed various hypotheses about how milky seas form. Investigators on the 1985 navy expedition theorized that the bioluminescent bacteria they collected had congregated around an algal bloom. Others have suggested that the steady glow results from “quorum sensing,” the ability of bacteria to communicate through chemical signaling. Once their density is high enough to produce a perceptible collective glow, they sustain a continual shine. But why?

Some biologists think bioluminescence in other marine organisms helps them attract food or mates or functions as a kind of burglar alarm, flashing when they are under attack in hopes of attracting the predators of their predators. The glow of quorum sensing in bacteria may act as a different invitation: When a colony runs low on food in the open water, it may glow to encourage nearby fish to come and consume the bacteria, consequently sustaining the bacteria in their guts.

The decade of DNB data has revealed when milky seas are most likely to occur. The peaks appear to be strongest in the northwestern Indian Ocean when winter and summer monsoons trigger phytoplankton blooms by bringing deep, cold, nutrient-rich water to the sea surface. Farther east, milky seas may be set up by the Indian Ocean Dipole, an El Niño–like pattern of sea-surface temperatures associated with cool, dry conditions and strong winds in the eastern Indian Ocean between May and October.

The satellite data also suggest an explanation for why the glow occasionally seems to extend to some depth, creating the perception among mariners that their ship is suddenly floating in light. Miller found that several milky seas occurred in the relative calm between large ocean eddies, where a combination of currents and temperature gradients can isolate a column of seawater from the surrounding ocean, putting it at a standstill. Such conditions, he hypothesized, could foster superdense bacterial populations whose quorum sensing extends vertically as well as horizontally to adjacent colonies, magnifying the depth and breadth of the resulting milky sea.

A Common Occurrence

Since most bioluminescent organisms live in the ocean, many at great depths, observing bioluminescence firsthand has required considerable resources—and risk. Marine biologist Edith Widder began her pioneering bioluminescence studies in the 1980s. She and other researchers who have taken deep-sea voyages have known for decades that bioluminescence is common.

The first reliable estimate of its occurrence came in 2017 from MBARI analysis of 17 years of video observations collected by remotely operated vehicles off the California coast. From more than 350,000 observations, which included more than 500 groups of organisms, Haddock and fellow researcher Séverine Martini concluded that at least three-quarters of the organisms were capable of bioluminescence.

The percentage remained remarkably consistent at different ocean depths. In a 2019 study, they found that about a third of the organisms living on the ocean floor are bioluminescent. Given that the ocean is the largest living space on the planet, the two analyses suggest that bioluminescence is one of the predominant ecological traits on Earth. Martini, Haddock, Widder, and the few other marine bioluminescence researchers are interested in its ecological functions, evolutionary history, chemistry, and genetics.



Ecological Importance

Humans have benefited greatly from bioluminescent species. Medical and biological researchers frequently use green fluorescent protein, which biologists isolated from bioluminescent jellyfish in the 1960s, as a visual marker of proteins and the components of living cells.

Widder is using bioluminescent bacteria to identify pollution hotspots in Florida's Indian River lagoon, one of the most diverse estuaries in North America. Fertilizer and pesticide runoff from farms and lawns, as well as leakage from sewage and septic systems, has been poisoning the lagoon for decades, and the pollution has accumulated in its sediments.

Because most pollutants interfere with bacterial respiration and therefore with bacterial bioluminescence, Widder and her colleagues have taken sediments from the lagoon and mixed them with bioluminescent bacteria in the lab to determine the relative concentrations of pollutants throughout the lagoon. This knowledge helps in monitoring, mitigation, and restoration efforts.

The rush to mine valuable metals from the ocean floor is a danger to bioluminescent organisms that communicate with one another across hundreds of feet of clear water. When robotic mining vehicles scrape the seafloor, they kick up clouds of sediment. After machines pump material to the surface and remove fist-sized, metal-rich nodules, they dump the remaining mud and silt back into the sea, again clouding water—inevitably disrupting communication among the bioluminescent organisms and interfering with their ability to find food and mates. The glowing seas that so terrified generations of mariners have taken no victims and left no traces; the cloudy seas created by humans, however, could permanently dim the ocean's light.

ABOUT THIS LECTURE

This content was adapted from the article "The Mystery of Milky Seas" by Michelle Nijhuis. It appeared in the August 2022 issue of *Scientific American*.

2



The True Colors of Dinosaurs

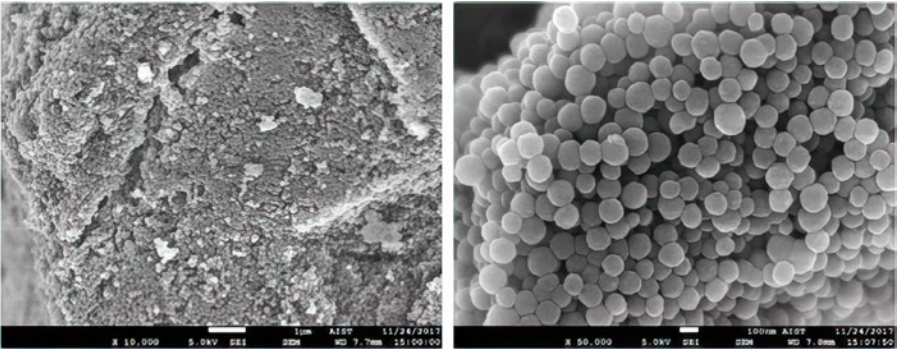


FOR DECADES, SCIENTISTS HAVE ASSUMED that pigments hardly ever survive the fossilization process. The few known examples all came from fossils of invertebrate creatures, not backboned ones. Researchers could only guess at the colors of most long-vanished animals, using modern ones as a guide. But discoveries made over the past dozen years are revealing ancient creatures' true colors. By studying the shapes and organizations of melanin-bearing structures, scientists have been able to deduce the actual colors and patterns of extinct dinosaurs and other animals from deep time. This lecture explores what scientists have learned about the physical appearances of certain creatures and how these clues have led to intriguing insights into their behaviors and habitats.



The Search for Fossilized Melanin

Jakob Vinther sat in a dark laboratory at Yale University in 2006 and zoomed into the fossilized ink of a 200-million-year-old squid relative under an electron microscope. He was riveted. These ancient structures looked exactly like the granules of melanin pigment that color the ink of modern squid and octopuses.



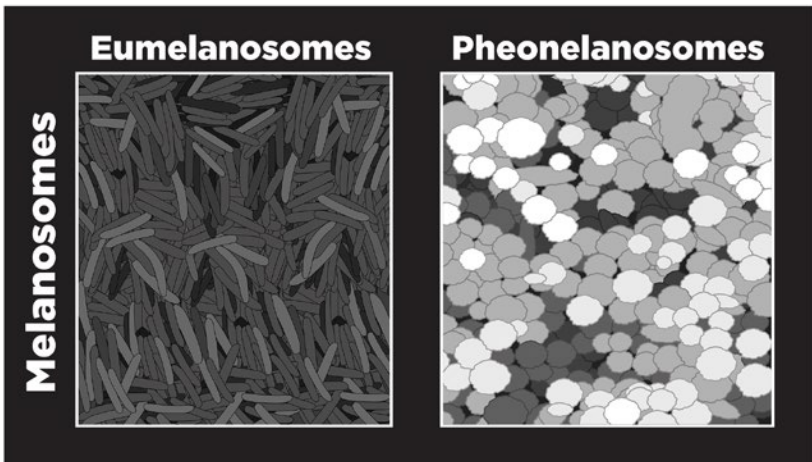
The consistently superb preservation of the ink made Vinther wonder whether melanin might persist in fossils of other kinds of organisms. Melanin is the same pigment found in hair, skin, feathers, and eyes. It can impart red, brown, gray, and black hues and create metallic sheens. Vinther, who is now a lecturer at England's University of Bristol, thought that if he could find melanin in other fossils, perhaps he could reconstruct the coloring of extinct animals, including dinosaurs.

Certain parts of organisms' bodies are likely to contain melanin: the outer covering of the body and the eyes. Those areas had to be examined under the electron microscope, which might require cutting a fossil specimen to get a sample. Well-preserved fossils are rare, however, and museums guard them closely.

Fortunately, a remarkable fossil site in Denmark had yielded exquisite bird fossils with feathers, which would be an ideal test case. Vinther convinced the curator of vertebrate fossils at the Geological Museum in Copenhagen to cut down a typewriter-sized block of limestone into a piece the size of a slice of bread so that it could fit into the museum's electron microscope. The stone contained a skull of a little bird with stains where the eyes used to be and a dark halo of feather impressions.

Melanosomes

Melanin is synthesized in specialized cells known as melanocytes by cellular components called melanosomes. Typically, the melanin remains encased in the melanosomes, which measure about 0.5 to 2 microns long and take two forms. There's a sausage-shaped kind that produces a form of melanin called eumelanin, which absorbs all wavelengths of light and thus gives squid ink and raven feathers their black color, and a meatball-shaped variant that makes pheomelanin, which creates a rusty red hue. An absence of pigments results in white plumage. Gray and brown colors, for their parts, appear to arise from combinations of eumelanin, pheomelanin, and pigment absence.



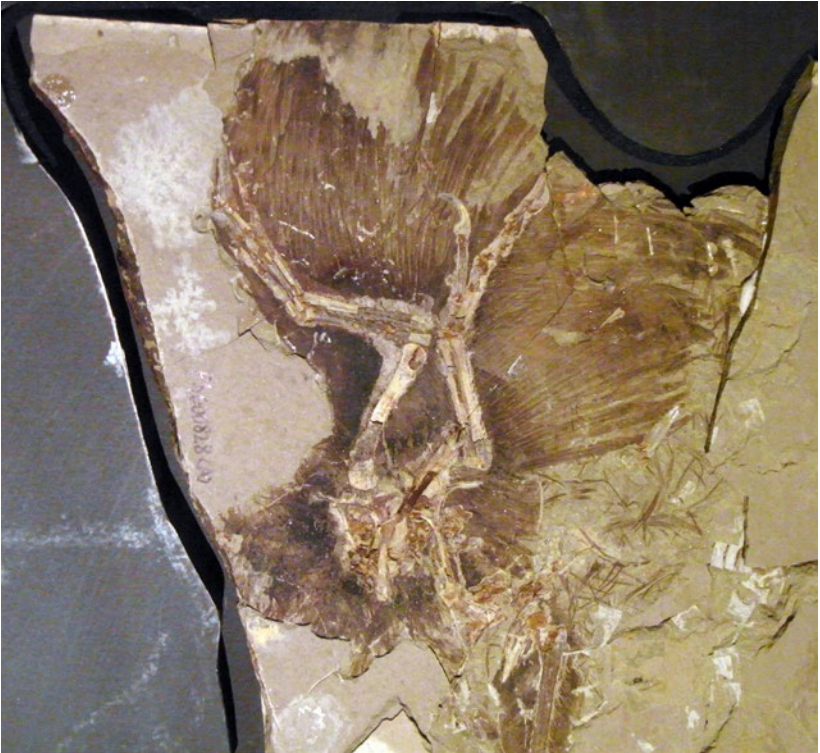
Yale's Richard Prum is one of the world's leading experts on bird color. He says the sausage-shaped melanosomes line up in a distinctive way along the barbs and barbules that constitute a feather's branches. The melanosomes arrive there during development, when the melanocytes transfer them into specialized cells called keratinocytes that give rise to feathers and hair. If the dark stains on the feather impressions evident in the Danish bird fossil came from melanin, then the microscope would reveal the sausages arranged this way along the feather branches.

When Vinther zoomed in on the fossil feathers, sure enough, he encountered millions of sausage-shaped structures. He immediately emailed the images to his PhD supervisor at Yale, Derek Briggs, a pioneer in the study of extraordinarily preserved fossils. But Briggs noted that these structures were the same as those he and others had found in fossil feathers and mammal hair for decades and had identified as bacteria. He was not convinced until he showed the images to Prum, who confirmed that they resembled melanosomes in every aspect.

Colors of Dinosaur Feathers

Scientists have since described melanosomes and other pigments from additional fossils. Some of the most spectacular findings have uncovered the colors of dinosaur feathers. Vinther, Matthew Shawkey, Liliana D'Alba, and others teamed up in 2009 to reconstruct the color pattern of *Anchiornis huxleyi*, a small, predatory, feathered dinosaur from China that lived around 155 million years ago. Like the Danish bird, the *Anchiornis* fossil had some dark stains visible to the naked eye, indicating the presence of organic material, probably melanin.

They had to develop a way to objectively predict colors from the shapes of the melanosomes. To do this, they studied melanosomes from 12 black, 12 brown, and 12 gray feathers of modern-day birds. By considering the length, width, and aspect ratio of the melanosomes, as well as how much they vary in shape, they could predict feather color using a statistical method called quadratic discriminant analysis with 90% accuracy.



When they applied this method to the melanosomes of *Anchiornis*, the results were striking. Their statistical predictions indicated that the feathers that covered much of the creature’s body were mostly gray. The long feathers on the animal’s arms and legs, in contrast, were unpigmented by melanosomes and thus white, except for the melanosome-laden tips, which they predicted were black. (Modern birds often have black-tipped wing feathers. The melanin, in addition to coloring the feathers, also strengthens them. Perhaps *Anchiornis* benefited from this property of melanin, too.) Most surprisingly, the feathers on the crown of the head contained impressions of round melanosomes—the “meatballs”—that would have given *Anchiornis* a ruddy crest. All told, this combination of colors made for a spectacular creature.

Iridescence

Since those early days, the feather data set has grown to hundreds of samples, including ones that allow scientists to accurately predict iridescence, the metallic sheen in the plumage of hummingbirds and peacocks, among other birds. Melanosomes responsible for this effect tend to be longer than typical melanosomes, and they may even be hollow or flattened. The iridescence arises from the packing of the melanosomes within the feather. Certain configurations of melanosomes refract light in ways that create different colors, depending on the angle at which the animal is viewed or illuminated.

Evidence of iridescence was found in a 49-million-year-old fossil feather from Messel, Germany, that preserves the original arrangement of melanosomes. They were packed into a dense, smooth layer found in the finest branches of the feather fossil, the barbules. The melanosomes occurred strictly on the farthest edge of the feather and on the top surface, the only part that was not obscured by other, overlapping feathers. That arrangement of melanosomes is known to produce what is called thin-film interference, the kind that occurs when gasoline floats on water and creates a rainbow of colors.

Soon, iridescence was found in a dinosaur—a crow-size creature from China with wings on all four limbs. Called *Microraptor*, it was a primitive cousin to *Jurassic Park*'s *Velociraptor*. The movie depicted *Velociraptor* with scaly skin, but scientists now know that both these dinosaurs were, in fact, covered in feathers. In *Microraptor*, the feathers preserve long, sausage-shaped melanosomes arranged to bend light in eye-catching ways. Its plumage thus would have been black, with the same sheen as a grackle's.

Colors Give Clues to Behaviors

Experts had presumed that *Microraptor* was nocturnal, based on the large size of its eye sockets. But the discovery that it possessed iridescent plumage suggests otherwise, because in modern birds, such coloration

is typically found in species that are active in the daytime. The bold coloring of *Anchiornis*, for its part, probably helped to attract mates or served as some other kind of display, as occurs in flashy modern birds.

Fossils of insects called lacewings offer a fascinating example of how pigments in one species can also reveal things about the other species it interacted with. Between 170 million and 150 million years ago, certain distinctive color patterns made their evolutionary debut in insects. Perhaps the most dramatic pattern to emerge during this time was the eyespot, a marking that resembles the eye of a different kind of animal and can startle predators, especially those approaching their prey at speed from a distance.

Lacewings are one of the first creatures known to have had eyespots. Their predators were most likely a small group of dinosaurs called the paravians, which are known to have lived at the same time as these lacewings and are thought to have given rise to birds. Although the fossil record of paravians themselves has been unable to pinpoint when flight evolved in this group, the appearance of these eyespots in the lacewings hints that some paravian dinosaurs had taken wing by this point and were exerting birdlike predation pressure on the insects.



Environmental Clues from Countershading

Other fossil melanosome discoveries have revealed details of the environment extinct organisms lived in. The fossil of a small, plant-eating dinosaur called *Psittacosaurus*, a relative of *Triceratops*, provided the first opportunity. These skeletons are common in northeastern China and are often complete.

In one specimen, a thin film drapes its body—the remains of the skin, including delicate scales. And its tail displays long, filamentous bristles that may be precursors to feathers. This fossil preserves evidence of beautiful color patterns all over the body. The patterns were subtle, with fine veining, dots, and stripes. And the animal had a dark back that gave way to a pale belly.

That kind of dark-to-light color gradient from back to belly counteracts the light-to-dark gradient created by illumination from the sun. This pattern, known as countershading, is common among modern animals ranging from dolphins to deer, helping both predators and prey blend in with their surroundings and thereby elude detection. The *Psittacosaurus* pattern provided the opportunity not only to study countershading in a dinosaur but also to deduce from the fossil alone what kind of environment the creature lived in and the light conditions under which it evolved its camouflage.

Innes Cuthill at the University of Bristol had recently studied countershading in modern ungulates: the group that includes horses, antelope, camels, pigs, and rhinoceroses. Because sunlight varies depending on the latitude at which an animal lives, as well as the density of vegetation in its habitat, the researchers had theorized that ungulate countershading, too, should differ according to latitude and habitat. Their findings bore that out.

Vinther's and Cuthill's teams knew from their visual inspection of the *Psittacosaurus* fossil that it had countershading of some sort. They therefore carefully projected the pigment pattern onto a life-size model of the dinosaur, which they accomplished by enlisting the help of

British paleoartist Bob Nicholls. Through this work, they determined that the transition from dark to light occurred low on the belly and tail in *Psittacosaurus*. Using another full-scale model and a variety of lighting conditions, they tested the function of the color pattern. They determined that the animal's coloring would have best camouflaged it in a habitat with diffuse light, such as that seen in a canopy forest.

A nearly six-meter-long ankylosaur, *Borealopelta*, was found in the approximately 112-million-year-old marine oil sand beds in northern Alberta, Canada. It was colored with reddish-brown pheomelanin and was countershaded. Living land animals of such size are not countershaded because there are no predators big enough to threaten them. In other words, for such a big creature to maintain its countershading from generation to generation, the Cretaceous predators must have been vicious enough to have threatened it. Perhaps that doesn't come as a big surprise considering the enormous, toothy theropod predators back then. But now, there's rock-solid proof in fossil color pattern of their gruesomeness.

As techniques advance, new discoveries will undoubtedly change our understanding of the past faster than ever before. Each one will bring us that much closer to seeing dinosaurs and other prehistoric creatures as they really were, in full Technicolor glory.

ABOUT THIS LECTURE

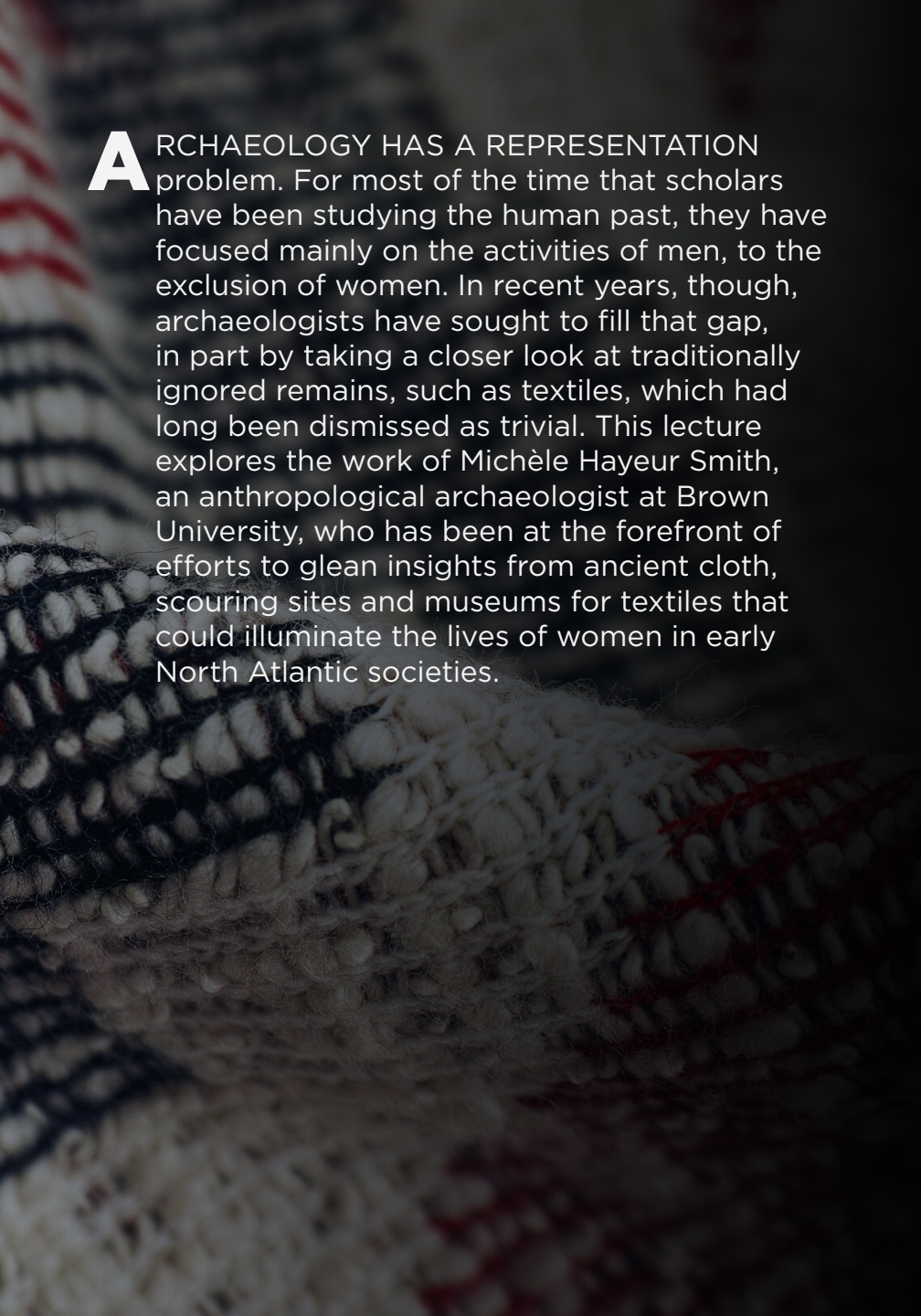
This content was adapted from the article "The True Colors of Dinosaurs" by Jakob Vinther. It appeared in the March 2023 collector's edition, *Revolutionary Science*.

A close-up photograph of a woven textile, likely a rug or blanket, featuring a complex pattern of red, white, and dark grey threads. The texture is highly detailed, showing the individual fibers and the tight weave.

3



How Textiles Gave Viking Women Power



ARCHAEOLOGY HAS A REPRESENTATION problem. For most of the time that scholars have been studying the human past, they have focused mainly on the activities of men, to the exclusion of women. In recent years, though, archaeologists have sought to fill that gap, in part by taking a closer look at traditionally ignored remains, such as textiles, which had long been dismissed as trivial. This lecture explores the work of Michèle Hayeur Smith, an anthropological archaeologist at Brown University, who has been at the forefront of efforts to glean insights from ancient cloth, scouring sites and museums for textiles that could illuminate the lives of women in early North Atlantic societies.



Scraps of Fabric That Tell a Story

Michèle Hayeur Smith's study of early North Atlantic textiles took off from the basement storage area of the National Museum of Iceland, where rows of metal shelving are bursting with boxes and bags of dirt-covered cloth. She first visited in 2009 to inspect the museum's collection of remains from the Viking Age and later periods. "It was literally thousands of fragments," she says. Yet they were just sitting there, hardly examined by anyone.

Hayeur Smith grew up surrounded by fabrics her anthropologist mother collected from around the world. In her twenties, Hayeur Smith earned a fashion degree in Paris. In the 1990s, as a PhD student at the University of Glasgow, she'd devoted herself to studying Viking women's dress and ornament, typically from artifacts found in burial sites. Inspired by her first glimpse of the wealth of textile remnants in the museum's storeroom, Hayeur Smith eventually decided to uncover the lives of the ordinary women who stood weaving at their looms.

Ever since then, she has been analyzing textiles spanning 900 years of history, starting with the Viking settlement of Iceland in the year 874. Her resulting studies of that museum's neglected collection of little brown scraps, as well as many other specimens of ancient Viking and later North Atlantic fabric, are among the first to prove the old guard wrong about the importance of cloth and women in ancient societies. In the Viking and medieval eras, women were the basis of the North Atlantic economy, and their cloth allowed people to survive the climate of the North Atlantic.

Portrayals of Viking Women

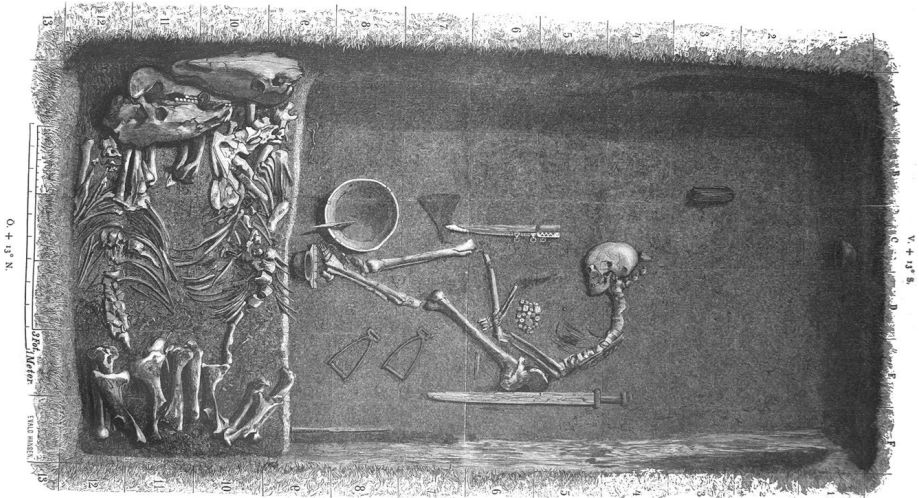
In popular culture, Viking women are seen through the eyes of the era. In the 1950s, they were portrayed as weak and subservient to men. In the 1970s, they were sexualized. In recent shows, such as *Vikings* and *The Last Kingdom*, they are depicted as shield-maidens or warriors.

Until Hayeur Smith began her work, the real lives of Viking women were largely unknown to science. According to archaeologist Douglas Bolender of the University of Massachusetts Boston, who studies the Viking Age and the medieval North Atlantic, the basic outline of Viking society came from the Icelandic sagas. Those book-length narrative accounts were set down more than 300 years after the events they describe. And the authors, who were men as far as we know, were Christianized people writing about their “pagan” ancestors.

Viking women have long been stereotyped in archaeology as performing primarily domestic tasks: child-rearing, cooking, weaving, and making clothing. Written accounts and archaeological evidence confirm that they were weavers. Yet for years at a time during their husbands’ absences for raids or trading expeditions, women ran the farms and engaged in trade, Hayeur Smith says.



Traditional views of women still color researchers' interpretations of evidence, says archaeologist Marianne Moen of the Museum of Cultural History in Oslo. She is a Viking expert who studies gender in the archaeological record, and she regularly sees how the meaning of artifacts is distorted by preconceptions of what they must signify. For example, a grave filled with a warrior's weapons at the Viking site of Birka in Sweden was long thought to be a man's final resting place until DNA evidence proved it was a woman's grave.



Weaving Homespun Cloth

There had been a few earlier studies of textiles, most notably by the late Elsa Guðjónsson, whose work was foundational for Hayeur Smith. Guðjónsson, however, had been able to study only “a handful” of archaeological pieces of fabric from the mountain of artifacts in the Iceland museum’s collection. And Guðjónsson’s work, like that of other textile analysts, focused mainly on technical details, such as thread counts, weave types, fleece varieties, embroidery stitches, and tools.

For Hayeur Smith, the technical details were important, but she had a different goal: to create what she calls a “social archaeology” of the culture through which she could uncover the lives of the women who created the cloth. For this purpose, she focused on the everyday “homespun”—or plain woolen fabric—made by ordinary women, who left no elaborate graves on their farms throughout the North Atlantic. Their only memorials are the textiles they wove on their warp-weighted looms.

A wood horizontal bar resting on two vertical ones holds the separate vertical warp threads, which are weighed down taut by volcanic stones of the kind that dot the shores of Iceland. Holding a heddle rod to separate the warp threads, the weaver draws the continuous horizontal weft thread in and out of one or more warp threads. By varying the number of warp threads, weavers could create common Icelandic patterns, mostly basic weaves known as twills and tabbies.

Before about the year 1000, the loom would have been set up in a *dyngja*, a weaving hut, says University at Albany archaeologist Kevin Smith, who is Hayeur Smith’s husband and has excavated similar structures in Iceland. These pit houses are dug down 1.5 to 3 feet deep, sometimes with turf walls above the pit and sometimes with wood walls that would have provided a space high enough for people to stand and work.

With a stone-framed hearth in one corner to offer warmth and light, these small buildings—no more than 9 by 15 feet—would have provided an intimate space fitting a loom and perhaps three women, spinning, weaving, and sharing stories.





Cloth as Currency and Commodity

Hayeur Smith made several trips to the museum's basement laboratory in 2010, examining specimen after specimen under a microscope. As she worked, she entered her data and extracted small samples for further analysis and testing, including a type of radiocarbon dating called accelerator mass spectrometry.

For the next several years, she studied fabric remains from museum repositories in Iceland, Greenland, the Faroes, Scotland, and Norway. Somewhere between the first and second year of this endless and “filthy” job, soil all over her fingers, Hayeur Smith had her eureka moment. “Viking Age textiles were colorful and varied,” she says, “but in medieval times, there is a complete shift into standardized cloth.”

Regular modern cloth, she explains, could vary from 75 to 300 warp threads, but in Iceland and only in Iceland, from the 12th through 17th centuries, every textile from every site fell into a tight range of 4 to 15 warp threads. In addition, the spin direction of the yarn shifted almost completely from clockwise warp and weft to counterclockwise in the weft in the 11th century. All these details are specifications for legal cloth, called *vaðmál*. As Hayeur Smith says: “Women were making the money!”

Iceland's budding economic system was based on Norway's. Certain commodities—cloth, cows, butter, grain—were legally assigned a value based on their equivalent value in silver. Toward the end of the Viking Age, however, homespun woolen cloth became much more important as a form of exchange in Iceland than in Norway. Scholars believe this shift may have resulted from such factors as a scarcity of silver after the Vikings stopped raiding, population growth, and the colony's burgeoning wool production.

“Although its value was still measured, in theory, against silver, this cloth ... came to be legally regulated as an exchange good in and of itself,” Hayeur Smith wrote in her 2020 book *The Valkyries' Loom: The Archaeology of Cloth Production and Female Power in the North Atlantic*. The name *vaðmál*, she explains, is a combination of the Old Norse words *vað* (“stuff” or “cloth”) and *mál* (“measure”), meaning “cloth measured to a standard.” Women made all of the *vaðmál*. It could be used to pay taxes and tithes, but it could also be traded or sold for making clothes and other necessities.

Women's Power as Cloth Makers

Scholars knew about *vaðmál* in an “abstract” sort of way, Hayeur Smith says, because it was precisely defined in the medieval law books. But the legal texts never mention the women weaving it, she points out. And nobody checked the cloth remains to see whether they conformed to the specifications in the legal texts.

In tandem with her textile analysis, she examined the legal texts. Through painstaking inspection, she confirmed that the cloth Icelandic women wove conformed exactly to these standards: a tweed, spun clockwise/counterclockwise, woven with 4 to 15 warp threads per centimeter. The cloth was also supposed to measure 2 ells in width and 6 ells in length (a little more than 1 yard wide and 3 yards long in modern measurements). That unit of cloth was equal to a certain weight of silver.

Hayeur Smith suggests that women either created the specifications themselves or collaborated with men to do it. “It’s not the men sitting there writing books ... because [men] didn’t go near weaving,” she says. She grounds this assertion partly on evidence from poetic and mythological sources, including the Icelandic sagas, which provide clues to deep-seated attitudes toward women and weaving in the Viking Age and beyond.

According to Karen Bek-Pedersen, an expert on female aspects of Viking religion at Aarhus University in Denmark, the *dyngja* can be seen as a space “charged with a feminine energy that reaches beyond the abilities of ordinary human women.” In the literary canon, she says, men who hang out there and gossip with the women are portrayed as cowards or villains and invariably come to a bad end. The fact that the *dyngja* was a space men shunned weighs heavily in Hayeur Smith’s appraisal of women’s power in cloth-making.

After Scandinavia turned Christian around the year 1000, looms were brought into the main living area of the *skáli*, the longhouse. Weaving would have been carried out in a separate area or room, with the taboos about this women’s craft most likely undiluted. These taboos became a critical factor in women’s power as their cloth turned into a major driver of the Icelandic economy.

Greenland’s Shift in Cloth Style

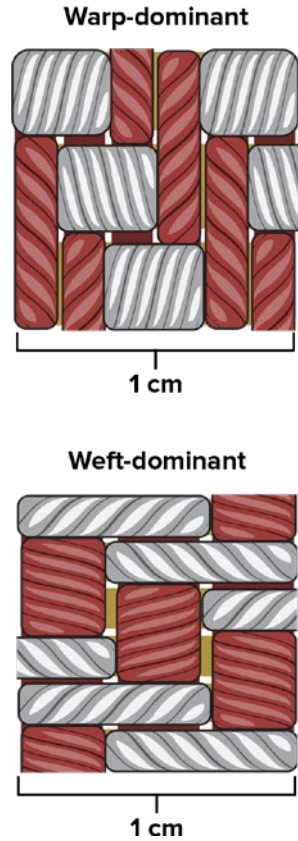
In 2011, Thomas McGovern, an archaeologist at the City University of New York, gave Hayeur Smith some fragile textile remains from an excavation he and his team had carried out at a site called Tatsipataa in southwestern Greenland. She was on a quest to find out why the cloth made by Greenland’s women diverged so much from the cloth made by Iceland’s female weavers.

Greenland had been settled in the year 986 entirely by Icelanders. The cloth there started out identical to the Icelanders' warp-dominant fabric but eventually shifted to contain more threads in its weft than its warp.

In the early 2000s, Else Østergård, a textile expert at the Danish National Museum, proposed an explanation for the shift. She thought it possible that Greenlandic women's weaving innovations might be a response to climate change during a period of cold climate known as the Little Ice Age. In Greenland, the first dramatic drop in temperatures started around 1340 and continued with fluctuations through the mid-15th century, when its colonies disappeared, to the 1900s.

Hayeur Smith set out to test Østergård's hypothesis against the archaeology, starting with the evidence from Tatsipataa. Collaborating with McGovern's doctoral student Konrad Smiarowski, she reviewed their excavation plan, a depiction of how the artifact layers were deposited over time. By dating the Tatsipataa cloth remains, she was able to correlate the ratio of weft to warp threads in each sample with published records of climate data. As Østergård had hypothesized, weft-dominant cloth did indeed increase as temperatures dipped in the 1300s.

But to prove that women were adapting their weaving to climate change, she had to collect remains from all over Greenland during different time periods. After inspecting some 700 cloth specimens from multiple archaeological sites across Greenland, she was able to track the evolution of the weft-dominant cloth. It becomes widespread between 1300 and 1362. "I confirmed [it was] climate change," she says.



Impact of Danish Rule on Women's Role as Weavers

Eventually, though, natural, political, and economic forces combined to strip Icelandic and Greenlandic women of the power they possessed as weavers of the all-important cloth. By around 1450, the Little Ice Age, among other factors, destroyed the Greenland Norse colony, and plague and political upheaval roiled the Kingdom of Norway.

In 1603, Danish authorities under King Christian IV imposed a royal monopoly on trading and strictly required all imports and exports to go through Denmark, reducing Iceland's freedom to trade. At this point, although Iceland continued to use *vaðmál* as currency and export it until the late 17th century, fish replaced the cloth as Iceland's primary export starting in the 14th century.

Imitating the English, whose male weaving guilds had produced fine cloths on foot-powered treadle looms since the 1300s, the Danes trained North Atlantic men to weave on these faster looms. They set up production workshops in locations around Iceland, including Reykjavik. They gave women spinning wheels, a much more efficient way to create yarn than the traditional spindle whorls used on the drop spindle.



The Danes also encouraged women to knit—a skill they’d learned in the 1500s—responding to a market demand for knitted exports. In addition, they imported fabric from Denmark to Iceland. Women could buy it to make clothing, saving them the relentless labor of weaving. In taking these measures, the Danes essentially pushed women out of the mainstream of weaving.

Nevertheless, as Hayeur Smith found in the archaeological record, women continued to weave their homespun cloth on their farms. She believes people used it as a statement of national identity in the face of Danish rule and the new laws imposing a transformation of the women’s 900-year-old tradition of textile production. She says: “I see it as resistance.” Yet the Danes—and the growth of industrialization—eventually prevailed. By the early 1800s, Hayeur Smith says, no one even knew how to weave on the old looms.

ABOUT THIS LECTURE

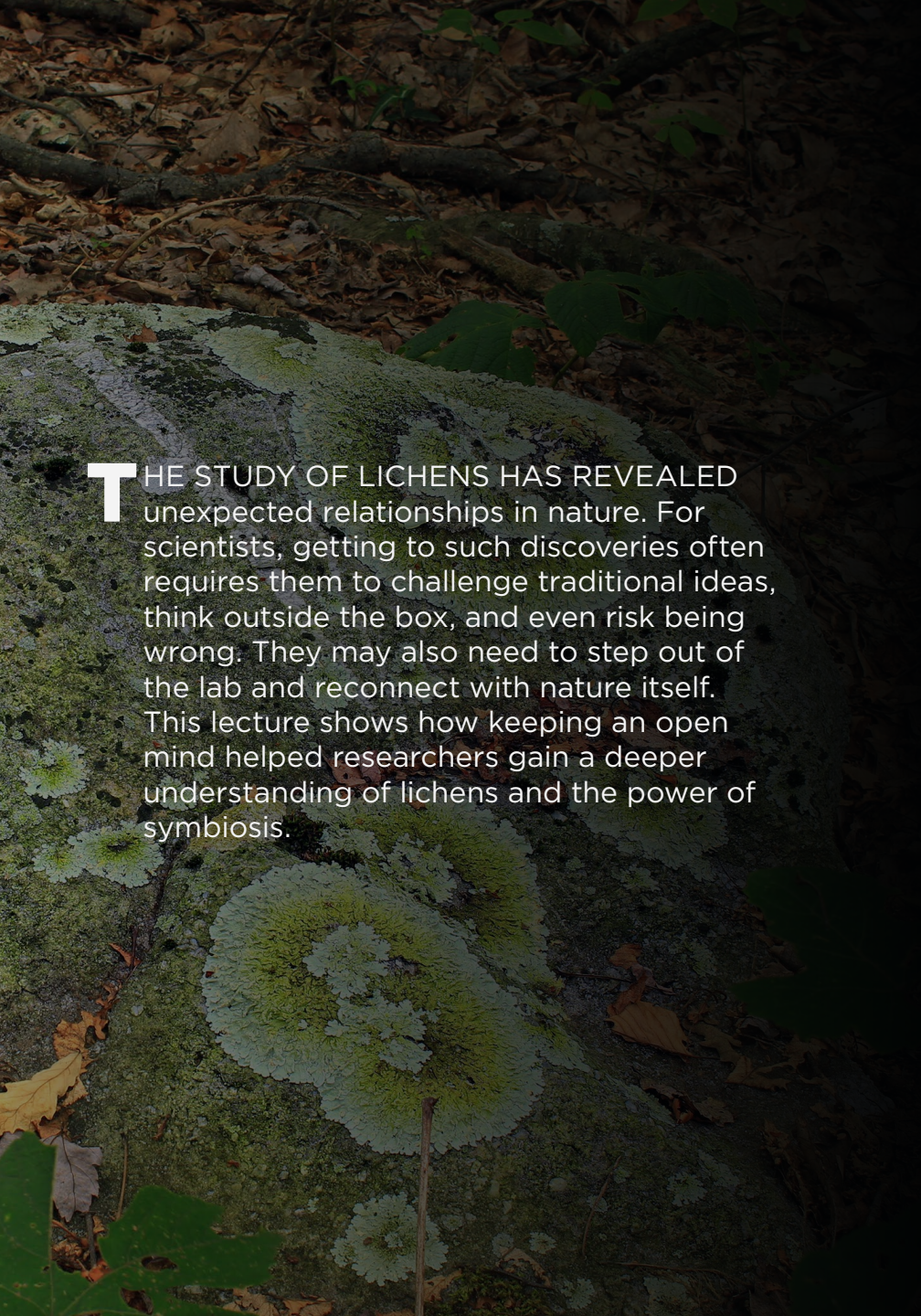
This content was adapted from the article “The Power of Viking Women” by Francine Russo. It appeared in the October 2022 issue of *Scientific American*.



4



New Truths about Lichens



THE STUDY OF LICHENS HAS REVEALED unexpected relationships in nature. For scientists, getting to such discoveries often requires them to challenge traditional ideas, think outside the box, and even risk being wrong. They may also need to step out of the lab and reconnect with nature itself. This lecture shows how keeping an open mind helped researchers gain a deeper understanding of lichens and the power of symbiosis.

A Maverick in the Scientific World

Trevor Goward spends much of his time walking along barely discernible deer and bear paths on his land adjacent to Wells Gray Provincial Park in British Columbia, a hand lens on a string around his neck, as an Australian shepherd named Purple trots along with him.

Goward's scientific love is lichens—those growths that look like little mosses or colored crusts stuck to trees and rocks everywhere. Wells Gray's 1.3 million acres were shaped by volcanoes and glaciers; its river valleys, sheer rock mountains, alpine meadows, and waterfall spray zones foster rich biodiversity.

Goward stumbled upon lichenology when he was educating himself about different branches of nature. Since then, despite being self-taught, he has become the go-to expert in central British Columbia for everyone from atmospheric scientists to gold prospectors to caribou biologists. Several lichen species have been named after him. He has published three taxonomic guides to lichens and has earned a spot as an associate member of the University of British Columbia botany department.



And yet Goward is a maverick in the scientific world. His radical thought experiments about lichens, published in 12 provocative essays, available on his website, *Ways of Enlichenment*, have been both ridiculed and lauded. They're largely ignored by most researchers because he holds no scientific degrees and because many of his ideas are not supported by rigorous data.

Still, Goward's astute observations and deep thinking follow in the footsteps of Charles Darwin's and Henry David Thoreau's approaches—which formed the basis of the theories of evolution and ecology. Toby Spribille, a lichenologist at the University of Alberta, says Goward's essays contain many gold nuggets and calls them “brilliant.”

Goward inspired Spribille's lab work while he was a postdoc at the University of Montana. That work paid off with a major advance: a 2016 cover story in the journal *Science* that rocked lichenology. The discovery called into question the very nature of the lichen symbiosis, shedding fresh light on how symbioses across biology work, how natural selection proceeds, and even how to define life-forms.

A Symbiotic Relationship

Lichens are ubiquitous. They are perhaps more than 500 million years old, occur on every continent, and can thrive in some of the most inhospitable places on Earth. They even survived for a year and a half in space, fully exposed to cosmic radiation, ultraviolet irradiation, and vacuum conditions.

The approximately 14,000 species of lichen come in a variety of forms: flat rounds on stones, scalloped leaves nestled among mosses, crusts clinging to tree bark, flowing strands hanging from branches, tiny trumpets tipped in red.

For centuries, people thought they were plants, and then fungi. Then, in the 1860s, Swiss botanist Simon Schwendener discovered that they were a partnership between a fungus, which is an organism classified in its

own kingdom because, unlike plants, it cannot make its own food, and an alga, an organism that feeds itself with photosynthesis but lacks the roots and stems of plants.

The fungus apparently provided the structure of the lichen, and the alga provided food for the fungus via photosynthesis. In some lichens, a cyanobacterium provides the food—and some types of lichen contain both an alga and a cyanobacterium, along with the fungus.

Schwendener's discovery was at first resisted by the scientific community, but it ultimately made lichens the textbook example of symbiosis, a mutually beneficial interaction among organisms. Since then, science has found symbioses across nature, including among the trillions of microbes that cling to our bodies.

A Third Symbiotic Partner

Science over the past two centuries has largely viewed molecules, cells, and species as individuals. Symbiosis challenges that notion. Just as light is both a wave and a particle, the fungus and alga are both individuals and parts of a whole. Spribille says science's reductionist focus has made it nearly impossible to fully understand symbiosis.

Spribille and his coauthors made their big reveal in 2016: Many lichens have a second fungus in the house. Their study began with a pair of lichens that Goward had drawn Spribille's attention to: *Bryoria fremontii*, which is hairlike and often brown and eaten by some indigenous people in the Northwest, and a similar lichen, *Bryoria tortuosa*, which is often a yellowish green and is toxic, with high levels of vulpinic acid.

The two posed a fascinating conundrum. Despite their differences, a genetic analysis by Saara Velmala of the University of Helsinki and her collaborators, including Goward, showed that the two species consisted of the same fungus and same alga. Spribille recalled how this finding perplexed them both.



Aside from their different appearances and levels of vulpinic acid, Goward observed that the two lichens also had slightly different ecologies. Although they grew in some of the same places, *B. tortuosa* was found only on the summer-dry fringes of *B. fremontii*'s larger territory.

He proposed that lichens are formed not by the shape of their fungal partner but by a series of decisions made during the developmental dance between fungus and alga. Goward suggested that the difference between the two species of *Bryoria* might stem from each of them having a different relationship with a third life-form, a bacterium.

After 5 years of work in the lab, Spribille and his colleagues discovered that both *Bryoria* species did include a third partner. But it was not a bacterium; it was another fungus, known as a basidiomycete yeast. The toxic *Bryoria* contained a lot more of the yeast than the edible one.

The team also demonstrated that the yeast was not a contaminant but had evolved with the other partners for more than 200 million years. Expanding their search to lichens across the globe, they found the yeast in 52 other sets, or genera, of lichen. The finding dramatically expanded the world's understanding of lichens, opening the door to other insights.

“Only now are we beginning to see that lichens really have pulled off a rare feat in evolution: a large multicellular organism but built entirely of microbes—and here’s the amazing thing—without a scaffold,” Spribille says. “Self-assembling, self-replicating, generation after symbiotic generation.”

Thinking outside the Box

Goward first became interested in *B. fremontii* and *B. tortuosa* when he read ethnobotanist Nancy Turner’s 1977 paper about *B. fremontii*’s importance to First Nations peoples. She said that women elders could easily distinguish the edible from the nonedible lichens. They used clues such as location, color, and the types of neighboring lichens to tell them apart.

When Stuart Crawford, a friend of Goward’s with a degree in ethnobotany, showed bundles of the two lichens to an elder and conservationist named Mary Thomas from the Neskonlith band, she correctly identified the edible one every time. Local people’s wisdom does not always agree with scientific explanations, Crawford says, but the result, based on observation, is correct.

Three months after the paper was published, Crawford told Spribille something amazing. For years, Crawford had been collecting writings from around the world—ancient Egypt, modern Mexico, medieval Russia, the biblical Middle East, a European cookbook from the 1950s—of people using lichens to make bread and alcoholic beverages. In some cases, they were using them explicitly for leavening and fermentation. On some level, Crawford realized, people knew that lichens contained yeast or functioned like yeast.

Spribille had a hard-won path to science. He grew up in a fundamentalist Christian family in northwestern Montana, where his parents pulled him out of school after fourth grade to protect him from “the influences of the world.” But his intellectual curiosity could not be contained. He eventually earned a PhD in lichenology at the University of Graz in Austria and in 2017 began his appointment as assistant professor of the ecology and evolution of symbiosis at the University of Alberta.

During his postdoc at the University of Montana, he met John McCutcheon, one of the coauthors on the *Science* paper and head of the lab in which the work was done. McCutcheon credits the breakthrough to technological advances that allowed the researchers to find the tiny yeast and to cooperation among diverse coworkers.

But also critical, he says, was Spribille's ability to look beyond what was assumed to be true. Spribille credits Goward with having "a huge influence" on his thinking. Goward's essays, he says, "gave me license to think about lichens in a way that was not orthodox and freed me to see the patterns I worked out in *Bryoria* with my coauthors."

Even so, Spribille says, "one of the most difficult things was allowing myself to have an open mind to the idea that 150 years of literature may have entirely missed the theoretical possibility that there would be more than one fungal partner in the lichen symbiosis." He warns that academia's emphasis on the canon of what others have established is inherently limiting.

Also worrisome to Spribille is that his own students are petrified of being wrong, a psychological state incompatible with breakthroughs. For a counterexample, he points to Goward. In the case of *Bryoria*, Goward surmised that a third partner was present, although he incorrectly thought it was a bacterium. But being correct "is not the criterion of a brilliant mind," Spribille says. On the contrary, brilliant minds are characterized by indefatigable curiosity and questioning, traits that Spribille tries to encourage in his students.

The Importance of Diverse Perspectives

Some of the most serious problems science is trying to solve today—climate change, loss of biodiversity, food and water insecurity—require big, integrated views from multiple perspectives. Stepping out of the lab and back into nature to observe how natural systems actually work is a critical first move. One biologist at the University of the South challenged himself to try it.

David George Haskell spent a year sitting in a square yard of old-growth forest in Tennessee, just observing, and wrote a Pulitzer-nominated book about it, *The Forest Unseen*. The experience was profoundly humbling. He says, “You wake up to the extent of your own ignorance.”

From that humility sprouted seeds of curiosity and dozens of questions about relationships among plants and animals, their ecological history, and how that related to climate and geology. Haskell is now an adviser to the New York City–based Open Space Institute, helping it to identify lands for conservation that are most likely to be climate-resilient.

If knowledge comes mostly through reading scientific literature, “we’re several steps removed from the actual phenomena we’re discovering,” Haskell explains. And while instruments are important to help scientists understand the world, he says: “Our bodies come preinstalled with all these amazing apps, and they connect directly into our consciousness. Through literally coming back to our senses, we can learn so much about the world.”

But for scientists who may not have a year, or 30, to spend contemplating the wild, collaborations or friendships with people outside the academy or from different disciplines can open space for new discoveries. With the scientific world often reluctant to publish him, Goward spreads his ideas one person at a time. His house, named Edgewood Blue, is on 10 acres abutting Wells Gray. He and his partner, Curtis Björk, a botanist, host an ongoing parade of biologists, aspiring naturalists, poets, geographers, ecologists, astrophysicists, and journalists who stay for a day or a week or longer in return for doing a bit of work at Edgewood. Both Spribille and Crawford have become regular guests.

Goward would like to create a more formal venue for learning, to increase “biological literacy” in the next generation, and has offered half of his land as a research center to Thompson Rivers University in nearby Kamloops, where he grew up. He also periodically invites people from various disciplines to meet for a few days of discourse.

Shifts in Thinking about Systems

And always, Goward comes back to one of his pet lines of inquiry: What are lichens, really? Are they organisms? Fungal greenhouses? Algal farmsteads? Ecosystems? Networks?

Goward's essays argue for seeing lichens not as their fungal or algal parts or even as ecosystems or organisms. Rather, they are all these things, biological systems encapsulated in a membrane: lichens as emergent property. After all, the lichens that were sent into space survived when their algae alone did not.

Thinking of lichens as systems fits with a larger shift in biology from viewing the fundamental unit of life as the individual to that of community or partnerships. As Haskell says, "Whether it is the microbiome within the human body or trees interacting with fungal partners belowground or lichens ... we're seeing that networked relationships are more fundamental and persist longer within biological systems than individuals do."

To Goward, lichens are the organisms that are most obviously about relationships. As such, they provide insights into all of life. Systems only hold together in the long term if the parts consider themselves integral to the whole and if the whole protects the parts, as lichens do. Goward says, "That's what's going wrong with us. As individuals, we're not concerned with the whole."

ABOUT THIS LECTURE

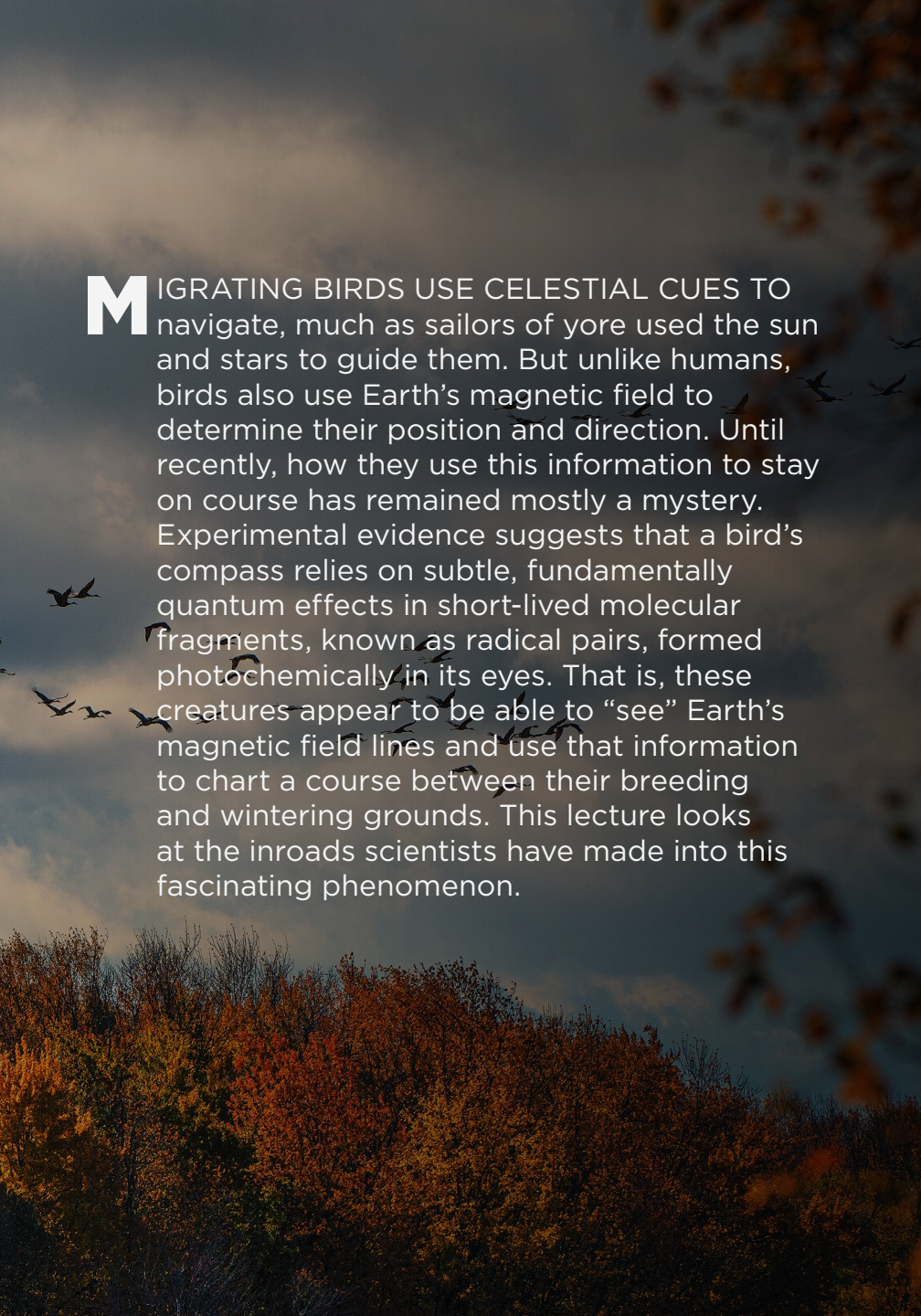
This content was adapted from the article "The Meaning of Lichen" by Erica Gies. It appeared in the March 2023 collector's edition, *Revolutionary Science*.

5



The Quantum Nature of Bird Migration



A flock of birds is seen in flight against a dark, cloudy sky. The bottom of the image shows a dense forest with trees displaying vibrant autumn foliage in shades of orange, yellow, and brown. The overall scene is atmospheric and suggests a natural setting.

MIGRATING BIRDS USE CELESTIAL CUES TO navigate, much as sailors of yore used the sun and stars to guide them. But unlike humans, birds also use Earth's magnetic field to determine their position and direction. Until recently, how they use this information to stay on course has remained mostly a mystery. Experimental evidence suggests that a bird's compass relies on subtle, fundamentally quantum effects in short-lived molecular fragments, known as radical pairs, formed photochemically in its eyes. That is, these creatures appear to be able to "see" Earth's magnetic field lines and use that information to chart a course between their breeding and wintering grounds. This lecture looks at the inroads scientists have made into this fascinating phenomenon.

Three Compasses

Every autumn, tens of thousands of bar-tailed godwits in Alaska complete one of the most impressive migrations on Earth: a nonstop transequatorial flight across the Pacific Ocean to New Zealand, 12,000 kilometers away. Billions of other young birds set out on similarly spectacular and dangerous migrations every spring, skillfully navigating the night skies without any help from more experienced birds. How do they find their way?

Migratory birds have an internal clock with an annual rhythm that tells them, among other things, when to migrate. They also inherit from their parents the directions in which they need to fly in the autumn and spring, and if the parents each have different genetically encoded directions, their offspring will end up with an intermediate direction. For example, if a southwest-migrating bird is crossed with a southeast-migrating bird, their offspring will head south when the time comes.



But how do the young birds know which direction is southwest or south or southeast? They have at least three different compasses at their disposal: One allows them to extract information from the position of the Sun in the sky, another uses the patterns of the stars at night, and the third is based on Earth's ever-present magnetic field.

In their first autumn, young birds follow inherited instructions, such as “fly southwest for 3 weeks and then south-southeast for 2 weeks.” If they make a mistake or are blown off course, they are generally unable to recover because they do not yet have a functioning map that would tell them where they are. This is one of the reasons why only 30% of small songbirds survive their first migrations to their wintering grounds and back again.

During its first migration, a bird builds up a map in its brain that, on subsequent journeys, will enable it to navigate with an ultimate precision of centimeters over thousands of kilometers. Equipped with this map, about 50% of adult songbirds make it back to their nesting site to breed every year.

Magnetoreception

Migratory birds' navigational input comes from several senses—mainly sight, smell, and magnetoreception. By observing the apparent nighttime rotation of the stars around the North Star, the birds learn to locate north before they embark on their first migration, and an internal 24-hour clock allows them to calibrate their sun compass. Characteristic smells can help birds recognize places they have visited before.

Scientists know a great deal about the detailed biophysical mechanisms of the birds' senses of sight and smell. But the inner workings of their magnetic compass have proved harder to understand.

The magnetic direction sense in small songbirds that migrate at night is remarkable in several important respects. First, observations of caged birds exposed to carefully controlled magnetic fields show that their compass does not behave like the magnetized needle in a ship's compass.

A bird detects the axis of the magnetic field and the angle it makes with Earth's surface, the so-called inclination compass. In laboratory experiments, inverting the magnetic field's direction so that it points in exactly the opposite direction has no effect on the bird's ability to orient correctly.

Second, a bird's perception of Earth's magnetic field can be disrupted by extraordinarily weak magnetic fields that reverse their direction several million times per second. Last, even though songbirds fly at night under the dim light of the stars, their magnetic compass is light-dependent, hinting at a link between vision and magnetic sensing.

Radical Pairs and Spin

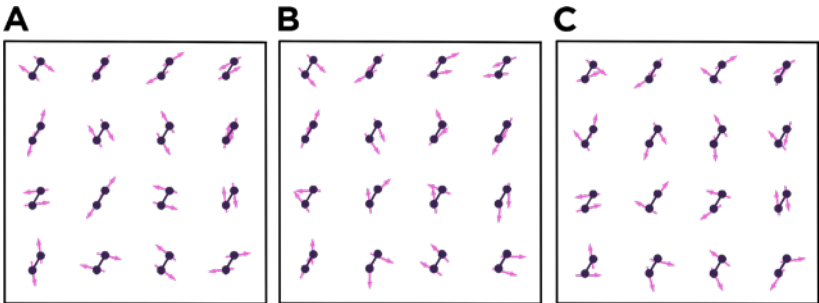
In 1978, in an attempt to make sense of these features of avian magnetoreception, Klaus Schulten, then at the Max Planck Institute for Biophysical Chemistry in Göttingen, Germany, put forth a remarkable idea: that the compass relies on magnetically sensitive chemical transformations.

At first glance, this proposal seems preposterous because the energy available from Earth's magnetic field is millions of times too small to break, or even significantly weaken, the bonds between atoms in molecules. But Schulten was inspired by the discovery 10 years previously that short-lived chemical intermediates known as radical pairs have unique properties that make their chemistry sensitive to feeble magnetic interactions.

To appreciate why radical pairs are so special, we need to talk about a quantum-mechanical property of the electron known as spin angular momentum, or "spin" for short. Spin is a vector with a direction as well as a magnitude, and it is often represented by an arrow, \uparrow or \downarrow , for example. Most molecules have an even number of electrons arranged in pairs with opposite spins ($\uparrow\downarrow$), which therefore cancel each other out.

Radicals are molecules that have lost or gained an electron, meaning that they contain an odd number and hence have a spin. Particles with spin act as microscopic magnets. When two radicals are created simultaneously by a chemical reaction (this is what we mean by *radical pair*), the two unpaired electrons, one in each radical, can have either opposite spins ($\uparrow\downarrow$) or parallel spins ($\uparrow\uparrow$).

Immediately after a radical pair is created, internal magnetic fields cause the two electronic spins to undergo a complex quantum “waltz” in which they flip between opposite and parallel states millions of times per second for periods of up to a few microseconds. Crucially, under the right conditions, this dance can be influenced by external magnetic fields.



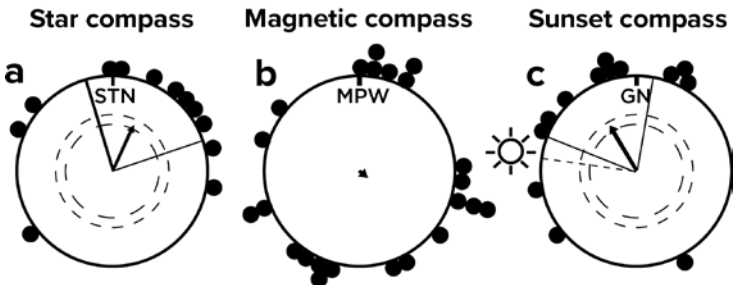
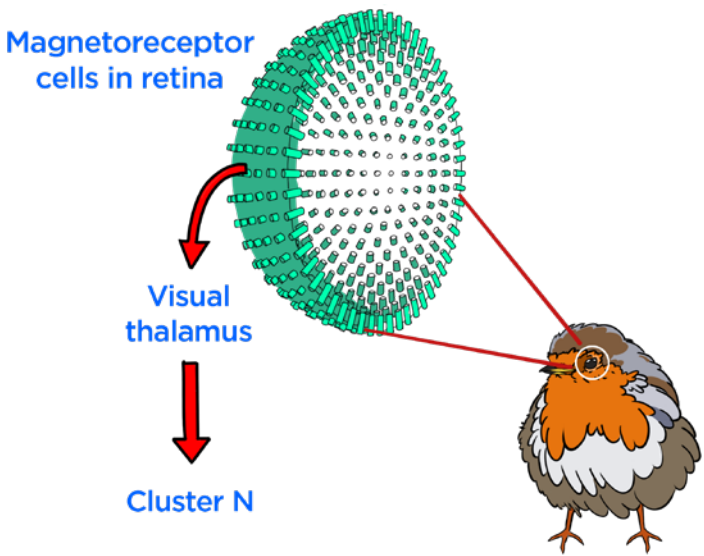
Schulten suggested that this subtle quantum effect could form the basis of a magnetic compass sense that might respond to environmental stimuli a million times weaker than would normally be thought possible. Research by Peter Hore, Henrik Mouritsen, and others has generated fresh support for this hypothesis.

Cluster N

To be useful, hypotheses need to explain known facts and make testable predictions. Two aspects of Schulten’s proposed compass mechanism are consistent with what is known about the birds’ compass: Radical pairs are indifferent to exact external magnetic field reversals, and radical

pairs are often formed when molecules absorb light. Given that the birds' magnetic compass is light-dependent, a prediction of Schulten's hypothesis is that their eyes play a part in the magnetic sensory system.

About 10 years ago, Mouritsen's research group at the University of Oldenburg in Germany found that a brain region called Cluster N, which receives and processes visual information, is by far the most active part of the brain when certain night-migrating birds are using their magnetic compass.



If Cluster N is dysfunctional, research in migratory European robins showed, the birds can still use their sun and star compasses, but they are incapable of orienting using Earth's magnetic field. From experiments such as these, it is clear that the magnetic compass sensors are located in the birds' retinas.

Cryptochromes

Schulten's hypothesis also predicts that there must be sensory molecules in the retina in which magnetically sensitive radical pairs can be created using the wavelengths birds need for their compass to operate, which another line of research had identified as light centered in the blue region of the spectrum. In 2000, he suggested that the necessary photochemistry could take place in a protein called cryptochrome.

Cryptochromes are found in plants, insects, fish, birds, and humans. They have a variety of functions, including light-dependent control of plant growth and regulation of circadian clocks. What makes them attractive as potential compass sensors is that they are the only known naturally occurring photoreceptors in any vertebrate that form radical pairs when they absorb blue light. Six types of cryptochromes have been found in the eyes of migratory birds, and no other type of candidate magnetoreceptor molecule has emerged in the past 20 years.

Like all other proteins, cryptochromes are composed of chains of amino acids folded up into complex three-dimensional structures. Buried deep in the center of many cryptochromes is a yellow molecule called flavin adenine dinucleotide (or FAD) that, unlike the rest of the protein, absorbs blue light. Embedded among the 500 or so amino acids that make up a typical cryptochrome is a roughly linear chain of three or four tryptophan amino acids stretching from the FAD out to the surface of the protein.

Immediately after the FAD absorbs a blue photon, an electron from the nearest tryptophan hops onto the flavin portion of the FAD. The first tryptophan then attracts an electron from the second tryptophan and so on. In this way, the tryptophan chain behaves like a molecular wire. The net result is a radical pair made of a negatively charged FAD radical in the center of the protein and, two nanometers away, a positively charged tryptophan radical at the surface of the protein.

Peter Hore, working with colleagues at Oxford in 2012, carried out experiments to test the suitability of cryptochrome as a magnetic sensor. The study used cryptochrome 1, a protein found in the plant *Arabidopsis thaliana*. Using short laser pulses to produce radical pairs inside the purified proteins, they found that they could fine-tune subsequent reactions by applying magnetic fields.

This was all very encouraging, but, of course, plants don't migrate. It was almost a decade before similar measurements could be made on a cryptochrome from a migratory bird.

Examining Cryptochromes in Migratory Birds

The first challenge was to decide which of the six bird cryptochromes to look at. Cryptochrome 4a was chosen, partly because it binds FAD much more strongly than do some of its siblings. Also, levels of it in migratory birds are higher during the spring and autumn migratory seasons than they are during winter and summer, when the birds do not migrate.

Computer simulations performed by Ilia Solov'yov in Oldenburg showed that European robin cryptochrome 4a has a chain of four tryptophans—one more than in the cryptochrome 1 from *Arabidopsis*. Naturally, scientists wondered whether the extended chain had evolved to optimize magnetic sensing in migratory birds.

The next challenge was to get large amounts of highly pure robin cryptochrome 4a. This was accomplished by Jingjing Xu, a PhD student in Mouritsen's lab. The samples were shipped to Oxford, where Christiane Timmel and her husband, Stuart Mackenzie, studied them using the sensitive laser-based techniques they had developed specifically for that purpose. Their research groups found that both the third and fourth tryptophan radicals at the end of the chain are magnetically sensitive when paired with the FAD radical.

Researchers suspect that the tryptophans work cooperatively for efficient magnetic sensing, biochemical signaling, and direction finding. They also speculate that the presence of the fourth tryptophan might enhance the initial steps of signal transduction, the process by which nerve impulses encoding the magnetic field direction are generated and ultimately sent along the optic nerve to the brain.

One more cryptochrome finding deserves mention here. Researchers compared robin cryptochrome 4a with the extremely similar cryptochrome 4a proteins from two nonmigratory birds, pigeons and chickens. The robin protein had the largest magnetic sensitivity, hinting that evolution might have optimized robin cryptochrome 4a for navigation.

Electrosmog Interference

Chemists and physicists in the 1980s showed that fluctuating magnetic fields alter the way radical-pair reactions respond to static magnetic fields. Their work predicted that a weak radio-frequency electromagnetic field, fluctuating with the same frequencies as the waltz of new radical pairs, might interfere with the birds' ability to use their magnetic compass. Thorsten Ritz of the University of California, Irvine, and his colleagues confirmed this prediction in 2004.

In 2007, Mouritsen began similar behavioral experiments in his lab in Oldenburg—with intriguingly different results. During the spring and fall, birds that travel between nesting and wintering grounds exhibit a behavior called *Zugunruhe*, or “migratory restlessness,” as if they are

anxious to get on their way. When caged, these birds usually use their magnetic compass to instinctively orient themselves in the direction in which they would fly in the wild.

Mouritsen found that European robins tested in wooden huts on his university's campus were unable to orient using their magnetic compass. He suspected that weak radio-frequency noise (sometimes called electrosmog) generated by electrical equipment in the nearby labs was interfering with the birds' magnetic compass.

He and his team then lined the huts with aluminum sheets to block the stray radio frequencies. On nights when the shields were grounded and functioned properly, the birds oriented well in Earth's magnetic field. On nights when the grounding was disconnected, the birds jumped in random directions.

When tested in an unshielded wooden shelter outside the city and well away from electrical equipment, the same birds had no trouble detecting the direction of the magnetic field. These results are significant on several fronts. If the radio-frequency fields affect the magnetic sensor, that provides compelling evidence that a radical-pair mechanism underpins the bird's magnetic compass.

Protecting Migratory Birds

One consequence of the enormous distances migratory birds travel is that they face more acute threats to their survival than most species that breed and overwinter in the same place. It is more difficult to protect them from the harmful effects of human activity, habitat destruction, and climate change. Relocating migratory birds away from damaged habitats is rarely successful because the birds tend to instinctively return to those unlivable locales.

The hope is that by providing new and more mechanistic insights into the ways in which these extraordinary navigators find their way, conservationists will have a better chance of “tricking” migrants into believing that a safer location really is their new home.


ABOUT THIS LECTURE

This content was adapted from the article "The Quantum Nature of Bird Migration" by Peter Hore and Henrik Mouritsen. It appeared in the April 2022 issue of *Scientific American*.

6



Unexpected Smashups in the Solar System



TEXTBOOKS OFTEN DESCRIBE THE formation of the solar system as occurring in a quiet, stately manner over hundreds of millions of years, with planets growing from small to large. But now, that tame sequence of linear growth is being replaced as discoveries suggest a more chaotic process occurred within a much shorter time frame. This lecture looks at how scientists are piecing together a new vision of planetary formation.

Rethinking Planetary Evolution

In 2009, a research team at the Massachusetts Institute of Technology (MIT) showed that a meteorite called Allende harbored signs of an ancient magnetic field in its rock. The discovery was a surprise. Allende had crashed into Earth in a huge fireball in Mexico in 1969 and contained some of the oldest known material in our solar system. It was supposed to be a fragment from a planetesimal—an early, nascent planet—that had been only slightly warm. Scientists presumed it had never gotten hot enough to melt the metal it contained. How, then, could this ancient piece of our solar system have become hot enough to create a magnetic dynamo?

Ben Weiss, a member of the team at MIT, brought up the problem with his colleague Linda T. Elkins-Tanton. Her students had just been asking her about planetary evolution, and she was rethinking some textbook wisdom, so she had the beginning of a new idea that might help answer the question.

Planetesimals had long been known to contain short-lived, unstable aluminum atoms radiating nuclear energy. This radioactive isotope is called aluminum-26, and when it decayed, that excess energy could have heated planetesimals. Conceivably, the heat from aluminum-26 in Allende's parent body could have climbed so high that the object actually melted from the inside out.

Metal within the body would have separated from silicate minerals in the rest of the rock and formed a liquid core that started to spin as the space rock rotated, creating a magnetic dynamo. Meanwhile, the outside of the planetesimal would have been chilled by the cold of space, and cold rock and dust from our solar system's primitive disk would have kept adding to this rind.



This idea that early building blocks of the solar system contained so much energy was not the story many of us learned in school. That creation, 4.567 billion years ago, was thought to be like a minuet. The gas and dust of a molecular cloud spun down into a disk around a growing young star, and the gas and dust coalesced into many little boulders, each one of which gradually grew to tens to hundreds of kilometers in diameter.

These planetesimals themselves then rammed together to form larger bodies, each perhaps the size of Mars, called planetary embryos. Only after that did the temperature in this space nursery pick up. These embryos, which in growing had gained gravity strong enough to begin clearing their orbits of surrounding debris, then collided together and grew into planets. Eventually, components in these planets separated into the familiar churning metal core and a silicate mantle—hot, wildly volcanic places, antithetical to life.

That's the old view. Now, that tame sequence of dust to boulders to planetesimals to embryos to planets is being replaced. Actual planetesimal formation, once surmised to take place over hundreds of millions of years, happened in just about 3 million years.

If the age of our solar system, in human time, is now a day, this growing up happened during the very first minute. More energy in the early small components means they did not have to wait and grow before developing different layers. Relatively minute planetesimals could host processes previously believed to be confined to planets, from melting and degassing to the creation of magnetic dynamos and volcanism.

And things in this system did not simply grow from small to large. Often, large things blew apart into small again. If planet-size spheres formed in those early years, through collisions of these smaller high-energy bodies, then glancing, hit-and-run blows among planetesimals sometimes stripped or destroyed them instead. Their debris could hit other spheres, enlarging them to planet size. Planets could be built, torn asunder, and rebuilt in just 10 million years or less.

Estimating Timelines

Scientists have developed instruments capable of measuring the elements that make up space rocks to parts per million or even less. Because we have a good idea of how long it takes radioactive elements to decay into their daughter elements, such measurements allow scientists to date when the planets and planetesimals that shed these fragments formed and changed.

Teams all over the world measured collections of meteorites and found that planetesimals formed within the first few million years after the dusty disk began to cool, that many of our terrestrial planets could have formed within the first 10 million years, and that even most of Earth probably formed and differentiated into a core and mantle by a few tens of millions of years.

With ever-improving telescopes, we can see young stars growing in other parts of our Milky Way galaxy, and in some cases, we can see the dust and gas disk from which a star and its planets grow. By estimating the ages of stars orbited by planets and comparing those measures with estimates for stars surrounded only by disks of dust and gas, researchers determined that these disks last, on average, only 3 million years.

Planetesimals, therefore, have just 3 million years to grow, on average. Any dust and gas not accreted onto rocks by then is lost into the star or out into space, and no more material for planet building is available. Considering that theorists used to think accretion took hundreds of millions of years, this is quite an acceleration!

More evidence for this timing comes from using the decay of radioactive elements like a clock, which ticks away at a steady rate as one element turns into another. Meteorites that have fallen to Earth contain those elements. Most are pieces of asteroids, which themselves are primitive remnants of planetesimals.

One radioactive isotope of the element hafnium prefers to stay in silicate minerals, such as those in Earth's mantle. But it decays to an isotope of tungsten, which combines readily with metals that make up planet cores. This decay happens on a regular schedule: Half the hafnium turns into tungsten in 9 million years.

This system gives us the timescale of planets and planetesimals: Metal core formation scavenges tungsten from the silicate mantle and carries it into the growing core. When scientists measure the ratios of hafnium to tungsten in meteorites, the amount of that tungsten isotope in them gives the time since core formation.

Such isotopic measurements of iron meteorites—many of which presumably hail from the metallic cores of planetesimals—show that their parent bodies formed within just 500,000 years of the first solids condensing from our protoplanetary disk of dust and gas. That time is less than 10 seconds into our putative 24-hour solar system. If iron meteorites are core fragments of planetesimals smashed apart by impacts, then planetesimals must have actually formed, melted, and created iron cores within that tiny period.

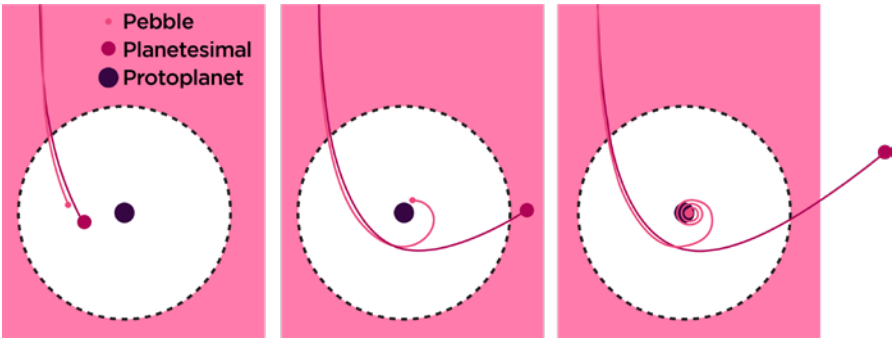
Investigating the Speedy Planet Growth

With experimental data firmly showing that the solar system formed much more quickly than textbook scenarios indicated, scientists who study it now had to explain how planets came into being so quickly.

Standard physics says that small clumps of dust that collide can readily stick together through electromagnetic forces, much as static electricity creates clumps of household dust. As the clumps grow, though, they reach what is called the meter barrier. Before they reach a meter in diameter, these growing boulders become too large to stick by electromagnetic forces and too small still to stick through gravitational attraction. Impacts even at very low speeds cause these conglomerates to destruct rather than accrete.

Several ideas for how growth beyond the meter barrier occurs have been proposed. Most hypotheses involve concentrating material in the protoplanetary disk through various types of turbulence that slam particles together. Such whirlpool-like forces might include phenomena called Kelvin-Helmholtz eddies, which develop between the gas and dust layer of the disk and could effectively crush regions of material together into larger bodies.

Much of that work has been pioneered by Anders Johansen, now at Lund University in Sweden. Hal Levison of the Southwest Research Institute and Johansen have separately worked on a newer model, called pebble accretion. Calculations indicate that even the smallest pieces of dust and clumps can be gravitationally diverted over several orbits to add to a growing planetesimal and can do so fast enough to build planetesimals early in the life of the solar system.



Investigating the Melting Pattern

No kind of crushing, however, could have caused the planetesimals to differentiate into mantles and cores. If planetesimals were first formed of primordial disk material in which metals and silicates were mixed, then only high temperatures and at least partial internal melting would allow the metal to sink to the interior and form a core.

This is where ideas about the radioactive aluminum come in. Because aluminum is one of the six most common elements in stony material (the others are silicon, magnesium, iron, oxygen, and calcium), aluminum-26, which has a half-life of approximately 700,000 years, could easily have heated at least some planetesimals to a melting temperature.

But what would have kept them from melting completely, considering that the new observations strongly indicate that some of these planetary seeds had unmelted rinds on the outside? Size is part of the answer. In the case of Allende's parent planetesimal, to reach melting, the rocky body would have needed to be massive enough so that its interior produced heat faster than its exterior radiated it away.

But the short half-life of aluminum-26 means this growth had to be quick. To retain enough heat for the melting pattern being posited, the Allende meteorite's parent planetesimal would have grown to 10 kilometers or more in radius within about 2 million years of the first solids in the solar system—and it may have grown to as much as 200 kilometers in radius.

It had been thought that planetesimals either melted entirely or remained primitive. But Weiss and Elkins-Tanton, the scientists at MIT, were suggesting a hybrid, where the most primitive material in the solar system encased a planetesimal that had melted inside—an outer rind as well as a melted core. This made sense because the Allende meteorite—with its record of a magnetic field caused by a heated interior—consists of unheated, primitive material alone. The only place it could come from would be a cool outer rind.

Allende's parent planetesimal retained this primitive unmelted surface rind because it was chilled by the coldness of space and because dust in the cool protoplanetary disk continued to attach to it over time. Unchanged by heating, the rind was able to maintain a record of the magnetic field produced by the inner part of the planetesimal's structure, the melted core and its magnetic dynamo.

If the earliest solar system was truly populated by hundreds or even thousands of differentiated planetesimals, zipping around and generating intense heat and magnetic dynamos like tiny Earths, it implied the entire infant system contained a great deal more heat than geologists had ever thought.

Visiting a Metal Asteroid

Scientists would like to learn how our own planet's structure and composition came into being out of this energetic, often chaotic environment. But we cannot get a good look at Earth's core because it is too deep and has pressures too high to sample directly.

But perhaps one particular asteroid, a remnant of an ancient planetesimal, might be a decent stand-in. In 2015, Elkins-Tanton and several colleagues began to design a space mission to explore this possibility. They wanted to explore a place that could prove or overturn hypotheses. The best destination, they had decided, was a world made out of metal: the metal asteroid Psyche.

Psyche is one of the largest asteroids, about 200 kilometers in diameter, and is located between Mars and Jupiter. All the physical measurements we have—from radio telescopes bouncing waves off the body—indicate that it consists almost solely of iron and nickel.

It looks like it is a stripped-naked planetesimal core, a last remnant of the hit-and-run collisions that disrupted bodies in the early solar system. The orientation of particles in Psyche, like tiny magnetic compass needles, might tell us whether it had a magnetic dynamo. There might also be some remnants of its rocky exterior that tell us what the deep mantle of planetesimals looked like. If there were surface impacts on the naked metal, the splashes might have produced sharp metal cliffs that froze before they could fall back to the surface.

In January 2017, NASA officially selected the Psyche mission for flight. It launched in October 2023. For the first time, humankind will visit a metal body, and scientists are eager to learn what it will tell us about the building of planets.


ABOUT THIS LECTURE

This content was adapted from the article “Solar System Smashup” by Linda T. Elkins-Tanton. It appeared in the March 2023 collector’s edition, *Revolutionary Science*.

7

A photograph of two cyclists riding on a paved road. The cyclist in the foreground is a woman wearing a black tank top, black shorts, and a red and white helmet. The cyclist in the background is a man wearing a black and red jersey and a black helmet. They are both leaning forward in a racing posture. The background features a large, rounded mountain under a clear blue sky.

**Why Exercise
Doesn't Help
You Lose
Weight**



RESearch in recent years has revealed some surprising insights into human metabolism. Data indicate that, contrary to received wisdom, humans tend to burn the same number of calories regardless of how physically active they are. Yet we have evolved to burn considerably more calories than our primate cousins do. This lecture looks at how these results help to explain two puzzles that might seem disparate at first but are, in fact, related: first, why exercise generally fails to aid weight loss and, second, how some of humanity's unique traits arose.

The Hadza

Researchers who are interested in human evolution and ecology often focus on energy expenditure because energy is central to everything in biology. Life is essentially a game of turning energy into kids, and every trait is tuned by natural selection to maximize the evolutionary return on each calorie spent.

Ideally, the study population lives in the same environments in which the species originally evolved, where the same ecological pressures that shaped its biology are still at work. That is difficult to achieve with human subjects because most people are divorced from the daily work of acquiring food from a wild landscape. For nearly all the past 2 million years, humans and our ancestors have been living and evolving as hunter-gatherers. Farming only got going about 10,000 years ago; industrialized cities and modern technology are only a few generations old.



Populations such as the Hadza, one of the last hunter-gatherer populations left in the world, are key to understanding how our bodies evolved and functioned before cows, cars, and computers. Hadza live off wild plants and animals in the dry savanna wilderness of northern Tanzania.

Life for the Hadza is physically demanding. Each morning, the women leave the grass huts of camp in small groups, some carrying infants on their back in a wrap, foraging for wild berries or other edibles. Men cover miles each day hunting with bows and arrows they make themselves. When game is scarce, they use simple hatchets to chop into tree limbs, often 40 feet up in the canopy, to harvest wild honey. Even the children contribute, hauling buckets of water back from the nearest watering hole, sometimes a mile or more from camp.

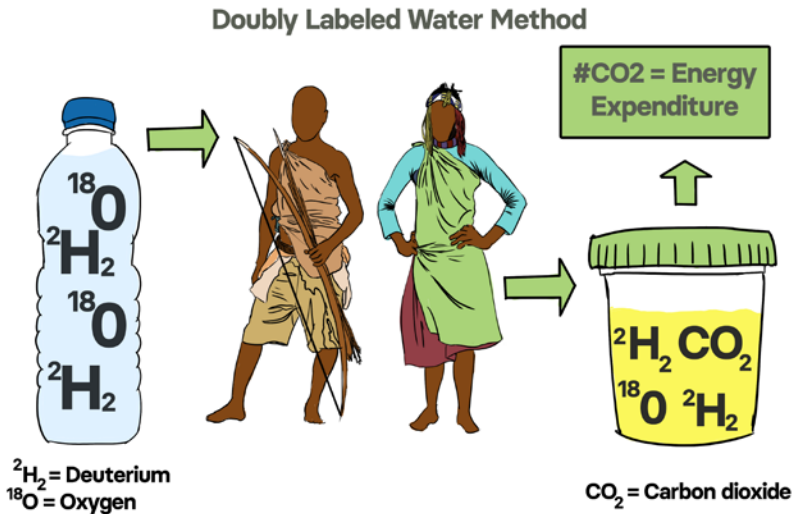
Hunting and gathering is cerebral and risky, a high-stakes game in which the currency is calories and going bust means death. Savvy is just as critical as stamina. Whereas other predators can rely on their speed and strength to obtain prey, humans have to outthink their quarry, considering their behavioral tendencies and scouring the landscape for signs of game.

Still, Hadza men land big game like giraffes only about once a month. They would starve if Hadza women were not executing an equally sophisticated, complementary strategy, using their encyclopedic knowledge of local plant life to bring home a reliable bounty every day. This complex, cooperative foraging is what made humans so incredibly successful and is the core of what makes us unique.

The Doubly Labeled Water Method

Researchers in public health and human evolution have long assumed that our hunter-gatherer ancestors burned more calories than people in cities and towns do today. Given how physically hard folks such as the Hadza work, it seems impossible to imagine otherwise. Many in public health go so far as to argue that this reduction in daily energy expenditure is behind the prevalence of obesity in the developed world, with all those unburned calories slowly accumulating as fat.

Anthropologists Herman Pontzer, Dave Raichlen, and Brian Wood set out to measure Hadza metabolism to determine the size of this energy shortfall and see just how deficient Westerners were in their daily expenditure. They enlisted a couple of dozen Hadza women and men to drink small, incredibly expensive bottles of water enriched in two rare isotopes, deuterium and oxygen-18. Analyzing the concentration of those isotopes in urine samples from each participant let the researchers calculate their body's daily rate of carbon dioxide production and thus their daily energy expenditure. This approach, known as the doubly labeled water method, is the gold standard in public health for measuring the calories burned each day during normal daily life.



To the researchers' surprise, when the analyses came back, the Hadza looked like everyone else. Hadza men ate and burned about 2,600 calories a day, Hadza women about 1,900 calories a day—the same as adults in the US or Europe. The team looked at the data every way imaginable, accounting for effects of body size, fat percentage, age, and sex. No difference. How was it possible?

It seems so obvious and inescapable that physically active people burn more calories that we accept this paradigm without much critical reflection or experimental evidence. But since the 1980s and 1990s, with the advent of the doubly labeled water method, the empirical data have often challenged the conventional wisdom in public health and nutrition.

The earliest doubly labeled water studies among traditional farmers in Guatemala, the Gambia, and Bolivia showed that their energy expenditures were broadly similar to those of city dwellers. In a study published in 2008, Amy Luke, a researcher in public health at Loyola University Chicago, took this work a step further, comparing energy expenditure and physical activity in rural Nigerian women with those measures in African American women in Chicago. Like the Hadza study, hers found no differences in daily energy expenditure between populations, despite large differences in activity levels.

The Body Adapts to Maintain Constancy

Humans are not the only species with a fixed rate of energy expenditure. On the heels of the Hadza study, Herman Pontzer piloted a large collaborative effort to measure daily energy expenditure among primates, the group of mammals that includes monkeys, apes, lemurs, and us. Pontzer's team found that captive primates living in labs and zoos expend the same number of calories each day as those in the wild, despite obvious differences in physical activity. Other studies have found similar energy expenditures in sheep and kangaroos kept penned or allowed to roam free, as well as in giant pandas in zoos and in the wild.

For a more granular look, comparing individuals instead of population averages, Pontzer, Luke, and Lara Dugas, also at Loyola, examined activity and energy expenditure in a large, multiyear analysis known as the Modeling the Epidemiological Transition Study (or METS).

More than 300 participants wore accelerometers, similar to a Fitbit or other fitness tracker, 24 hours a day for an entire week, while their daily energy expenditure was measured with doubly labeled water. The team found that daily physical activity, tracked by the accelerometers,

was only weakly related to metabolism. On average, couch potatoes tended to spend about 200 fewer calories each day than people who were moderately active: the kind of folks who get some exercise during the week and make a point to take the stairs.

But more important, energy expenditure plateaued at higher activity levels: People with the most intensely active daily lives burned the same number of calories each day as those with moderately active lives. How does the body adjust to higher activity levels to keep daily energy expenditure in check?

It could be that people with high activity levels change their behavior in subtle ways that save energy, like sitting rather than standing or sleeping more soundly. But analysis of the METS data suggests that although these behavioral changes might contribute, they are not sufficient to account for the constancy seen in daily energy expenditure.

Another intriguing possibility is that the body makes room for the cost of additional activity by reducing the calories spent on the many unseen tasks that take up most of our daily energy budget: the housekeeping work that our cells and organs do to keep us alive. For example, exercise often reduces inflammatory activity that the immune system mounts, as well as levels of reproductive hormones such as estrogen. In lab animals, increased daily exercise has no effect on daily energy expenditure but instead results in fewer ovulatory cycles and slower tissue repair. Humans and other creatures seem to have several evolved strategies for keeping daily energy expenditure constrained.

The Different Roles of Diet and Exercise

If daily energy expenditure has not changed over the course of human history, the primary cause of widespread obesity must be the calories consumed. This should not be news. Experts know from personal experience and lots of data that just hitting the gym to lose weight is frustratingly ineffective. But the new science helps to explain why exercise is such a poor tool for weight loss. It is not that we are not trying hard enough. Our bodies have been plotting against us from the start.

You still have to exercise. Exercise has tons of well-documented benefits, from increased heart and immune system health to improved brain function and healthier aging. In fact, metabolic adaptation to activity may be one of the reasons exercise keeps us healthy, diverting energy away from activities, such as inflammation, that have negative consequences if they go on too long.

The foods we eat certainly affect our health, and exercise paired with dietary changes can help keep off unwanted pounds once a healthy weight has been reached, but evidence indicates that it is best to think of diet and exercise as different tools with different strengths. Exercise to stay healthy and vital; focus on diet to manage your weight.

Expensive Traits Require a Faster Metabolism

A constrained, adaptive metabolism leaves researchers with larger, existential questions. If daily energy expenditure is virtually immobile, how could humans evolve to be so radically different from our ape relatives? Nothing in life is free. Resources are limited, and investing more in one trait inevitably means investing less in another.

It is no coincidence that rabbits reproduce prodigiously but die young; all that energy plowed into offspring means less for bodily maintenance and longevity. *Tyrannosaurus rex* can thank its big head of nasty teeth and powerful hind limbs for its puny arms and hands. Even dinosaurs couldn't have it all.

Humans flout this bedrock evolutionary principle of austerity. Our brains are so large that, as you read this lecture, the oxygen from every fourth breath you take is needed just to feed your brain. Yet humans have bigger babies, reproduce more often, live longer, and are more physically active than any of our ape relatives.

Our energetic extravagance presents an evolutionary puzzle. Primates burn only half as many calories a day as other mammals do. The reduced metabolic rates of primates correspond with their slow rates of growth

and reproduction. Perhaps, conversely, the faster reproduction and other expensive traits of humans were linked to the evolution of an increased metabolic rate.

All that was needed to test this idea was getting a bunch of frenetic chimpanzees, wily bonobos, phlegmatic orangutans, and skittish silverback gorillas to carefully drink doses of doubly labeled water without spilling and to provide a few urine samples. In a scientific tour de force, Steve Ross and Mary Brown, both at Lincoln Park Zoo in Chicago, worked with caretakers and veterinarians from more than a dozen zoos across the US to pull that off. It took a couple of years, but they accumulated enough data on great ape energy expenditure to provide a solid comparison with humans.

Sure enough, humans burn more calories each day than any of our great ape relatives. Even after accounting for effects of body size, activity level, and other factors, humans consume and expend about 400 more calories a day than chimpanzees and bonobos do; differences with gorillas and orangutans are larger still.

Those extra calories represent the extra work our bodies do to support larger brains, produce more babies, and maintain our bodies so we live longer. It is not simply that we eat more than other apes, although we do that, too. Our bodies, right down to the cellular level, have evolved to burn energy faster and get more done than our ape relatives.

A Faster Metabolism Makes Humans Unique

But here's the fundamental danger in the high-energy human strategy: Coming home empty-handed was both more likely and more consequential. Many of the energy-rich foods we need to fuel our faster metabolisms are inherently difficult to obtain in the wild, increasing the energy cost of finding food and heightening the risk of starvation for the Hadza men and women out foraging and their kids back at camp.

Happily for us, humans have evolved a few tricks to keep starvation at bay. We are the only species that has learned to cook, which increases the caloric value of many foods and makes them more efficient to digest.

We have also evolved to be fat. Even Hadza adults, lean by any human standard, carry twice as much fat as chimpanzees idling away in zoos. Our propensity to store fat most likely coevolved with our faster metabolism as a critical energy buffer to survive lean times.

And sharing food is so fundamental to the human experience that we take it for granted, but it is a unique and essential part of our evolutionary inheritance. Other apes do not share like we do.

Beyond our nutritional requirements and fixation with fat, perhaps the most profound impact of our increased energy expenditure is this human imperative to work together. Evolving a faster metabolism bound our fortunes to one another, requiring that we cooperate or die.

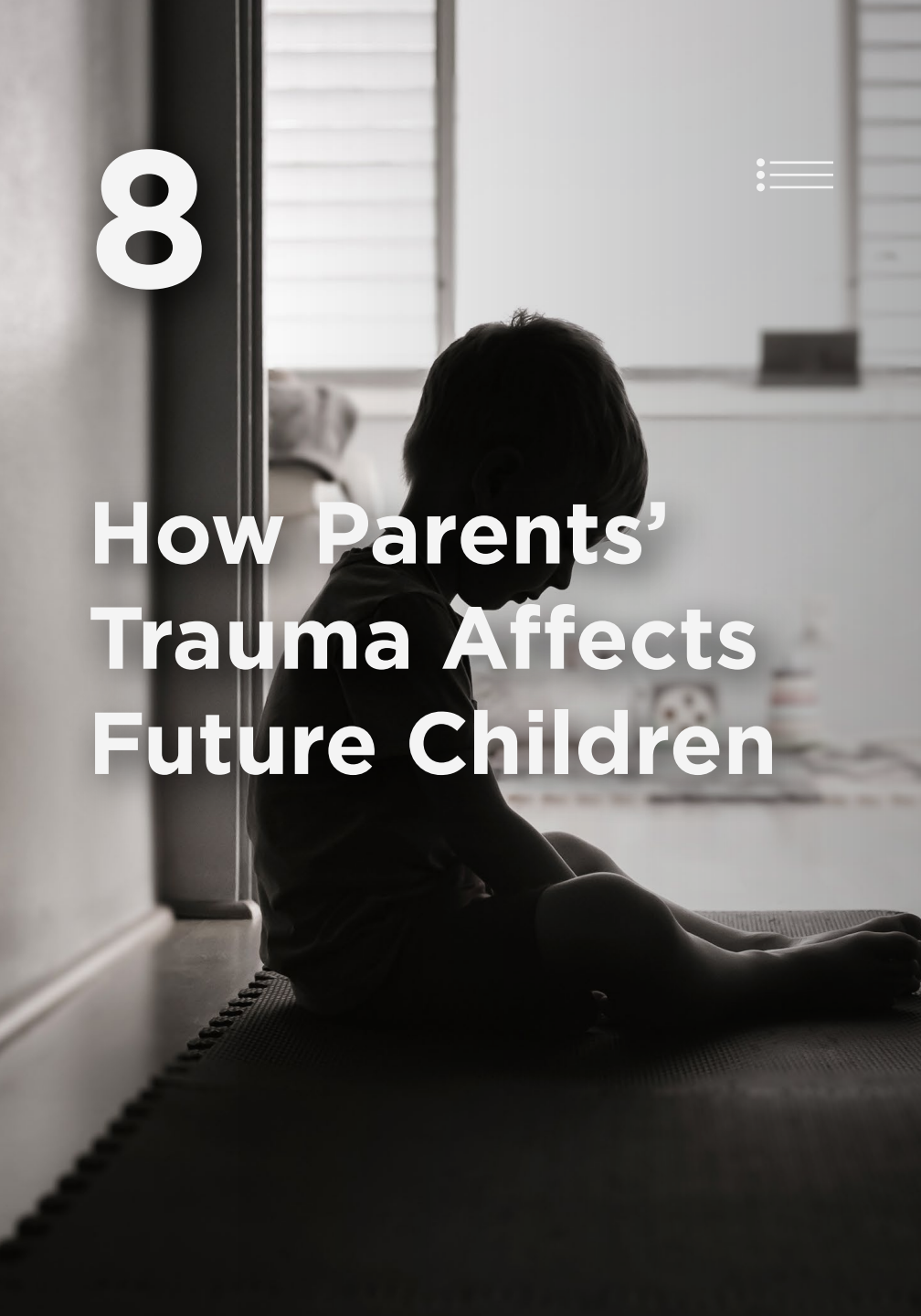
ABOUT THIS LECTURE

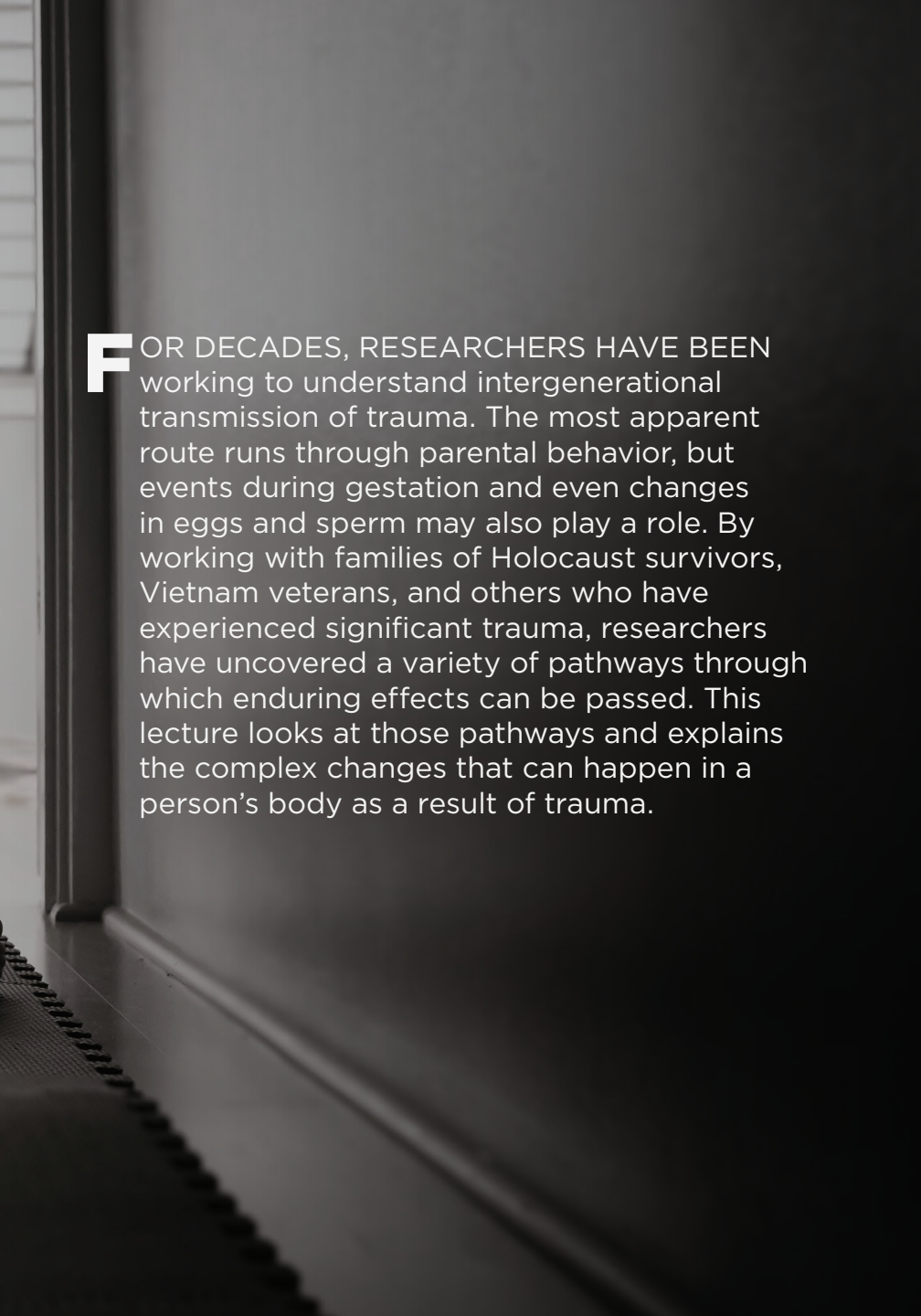
This content was adapted from the article “The Exercise Paradox” by Herman Pontzer. It appeared in the March 2023 collector’s edition, *Revolutionary Science*.

8



How Parents' Trauma Affects Future Children



A dark, moody photograph of a window frame. The window is on the left side, showing a view of a building exterior with a grid pattern. The rest of the image is a dark, almost black, gradient background. The text is overlaid on the right side of the image.

FOR DECADES, RESEARCHERS HAVE BEEN working to understand intergenerational transmission of trauma. The most apparent route runs through parental behavior, but events during gestation and even changes in eggs and sperm may also play a role. By working with families of Holocaust survivors, Vietnam veterans, and others who have experienced significant trauma, researchers have uncovered a variety of pathways through which enduring effects can be passed. This lecture looks at those pathways and explains the complex changes that can happen in a person's body as a result of trauma.

PTSD in Pregnant Women after 9/11

After the twin towers of the World Trade Center collapsed on September 11, 2001, in a haze of horror and smoke, clinicians at the Icahn School of Medicine at Mount Sinai in Manhattan offered to check anyone who'd been in the area for exposure to toxins.

Among those who came in for evaluation were 187 pregnant women. Many were in shock and at risk of developing post-traumatic stress disorder, or PTSD—experiencing flashbacks, nightmares, emotional numbness, or other psychiatric symptoms for years afterward. Were their fetuses also at risk?

When the women's babies were born, they were smaller than usual—the first sign that the trauma of the attack had reached the womb. Nine months later, when 38 women and their infants went for a wellness visit, psychological evaluations revealed that many of the mothers had developed PTSD. And those with PTSD had unusually low levels of the stress-related hormone cortisol, which was a feature that researchers were coming to associate with the disorder.

Surprisingly and disturbingly, the saliva of the 9-month-old babies of the women with PTSD also showed low cortisol. The effect was most prominent in babies whose mothers had been in their third trimester on that fateful day.

Just a year earlier, a research team had reported low cortisol levels in adult children of Holocaust survivors, but they'd assumed that it had something to do with being raised by parents who were suffering from the long-term emotional consequences of severe trauma. Now, it looked like trauma could leave a trace in offspring even before they are born.



Holocaust Survivors and Their Children

Intergenerational transmission of trauma came into the spotlight in the 1990s, when researchers led by Rachel Yehuda, a professor of neuroscience and psychiatry at the Icahn School of Medicine at Mount Sinai, documented high rates of PTSD among Holocaust survivors in her childhood community in Cleveland. The study was the first of its kind, and it got a lot of publicity; within weeks, a Holocaust research center was created at Mount Sinai. Many people called in to learn more. Some were Holocaust survivors, but most were the adult children of Holocaust survivors.

One particularly persistent caller, Joseph, insisted that researchers study people like him. "I'm a casualty of the Holocaust," he told Yehuda. Joseph was a handsome and successful investment banker. But he lived each day with a vague sense that something terrible was going to happen and that he might need to flee or fight for his life.

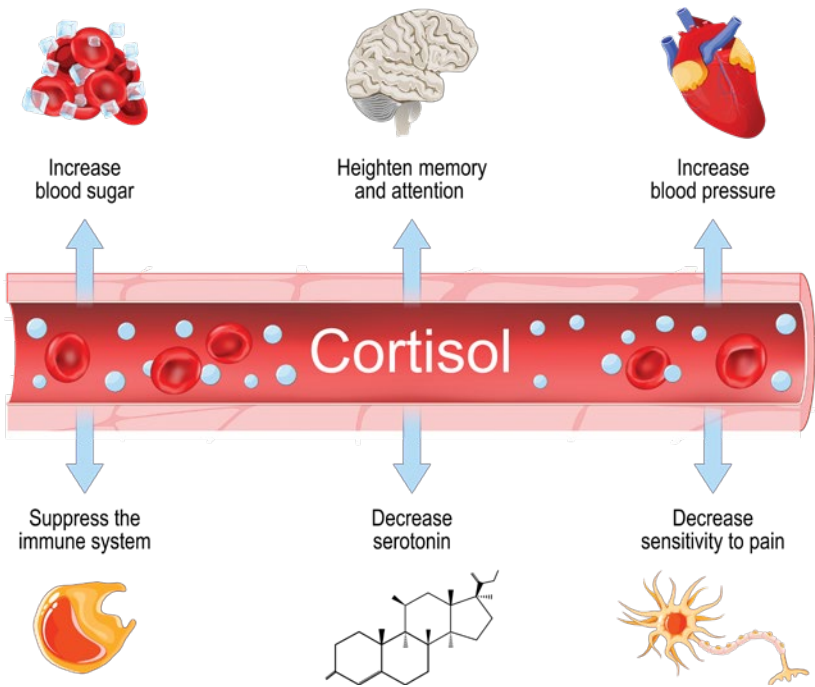
Joseph's parents had met in a displaced-persons camp after surviving several years at Auschwitz, then arrived penniless in the US. His father worked 14 hours a day and said very little, never mentioning the war. But almost every night, he woke the family with shrieks of terror from his nightmares. His mother spoke endlessly about the war, telling vivid bedtime stories about how relatives had been murdered before her eyes.

Joseph was one of many adult children of Holocaust survivors who suffered from anxiety, grief, guilt, dysfunctional relationships, and intrusions of Holocaust-related imagery. When Yehuda and other researchers evaluated the offspring of the Holocaust survivors they had just studied in Cleveland, the results were clear.

Survivors' adult children were more likely than others to have mood and anxiety disorders, as well as PTSD. Many Holocaust offspring also had low cortisol levels—something that we had observed in their parents with PTSD. What did it all mean?

The Role of Stress Hormones

Unraveling the tangle of trauma, cortisol, and PTSD has occupied Yehuda and other researchers for the decades since. In the classic fight-or-flight response, identified in the 1920s, a threat triggers the release of stress hormones, such as adrenaline and cortisol. The hormones prompt a cascade of changes, such as quickening the pulse and sharpening the senses, to enable the threatened person or animal to focus on and react to the immediate danger. These acute effects were believed to dissipate once the danger receded.



In 1980, however, psychiatrists and advocates for Vietnam War veterans won a prolonged struggle to get post-traumatic stress included in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-III)*. It was the first official recognition that trauma could have long-lasting effects.

But the diagnosis was controversial. Many psychologists believed that its inclusion in DSM-III had been politically, rather than scientifically, driven—in part because there were no scientific explanations for how a threat could continue to influence the body long after it was removed.

Complicating matters, studies of Vietnam veterans were generating perplexing results. In the mid-1980s, neuroscientists at Yale University reported that veterans with PTSD had high levels of adrenaline but lower levels of cortisol than patients with other psychiatric diagnoses.

Because stress usually causes stress hormones, including cortisol, to rise, many researchers, including Yehuda, were skeptical of these observations. But when a different group of veterans was studied, using other methods for measuring cortisol, the finding was replicated.

The Link between Low Cortisol and PTSD

For Yehuda's study demonstrating PTSD in Holocaust survivors, she and five researchers interviewed people and tested blood and urine samples. The results showed that half the Holocaust survivors had PTSD, and those with PTSD had low cortisol. There was no question about it—even if the traumatic experience happened long ago, PTSD went hand in hand with low cortisol.

But why? An important clue came from a 1984 review by the late Allan Munck and other researchers at the Geisel School of Medicine at Dartmouth. They noted that among stress hormones, cortisol played a special, regulatory role.

High levels of stress hormones, if sustained for a long time, harm the body in multiple ways, weakening the immune system and increasing susceptibility to problems such as hypertension. But in a context of acute trauma, cortisol may paradoxically also have a protective effect. It shuts down the release of stress hormones—including itself—reducing the potential damage to organs and the brain. A trauma-induced feedback loop could conceivably reset the cortisol “thermostat” to a lower level.

In the early 1990s, it had been shown that Vietnam veterans were more likely to develop PTSD if they'd been abused as children. Slowly, a connection was emerging between intense childhood adversity with low cortisol and the possibility of future PTSD.

If someone with low cortisol was subjected to a traumatic experience, the cortisol levels in their bodies might be too low to tamp down the stress reaction. Adrenaline levels might then shoot way up, searing the memory of the new trauma into the brain. Perhaps low cortisol marked a vulnerability to developing PTSD.

But what mechanism connected trauma exposure to low cortisol to future PTSD? It turned out that Vietnam veterans with PTSD had a greater number of glucocorticoid receptors. These are proteins that cortisol binds to in order to exert its diverse influences.

That suggested a greater sensitivity to cortisol: A small increase in the hormone's concentration would cause a disproportionate physiological reaction. When scientists looked more closely at the molecular underpinnings of cortisol functioning, they began to understand how exposure to trauma might reset the cortisol feedback loop.

Epigenetic Changes on Stress-Related Genes

In the 1990s, scientists were realizing that the products of our genes are sensitive to factors that are not written directly into our genetic code. Genes provide the templates for producing proteins, but how much of

those proteins gets produced depends on the environment. The discovery gave rise to epigenetics, the study of what influences gene expression. It proved crucial for understanding both the neurobiology of PTSD and the intergenerational effects of trauma.

Epigeneticists explore the switches that turn gene expression on and off. One such mechanism, called methylation, is a process by which methyl groups attach to key sites on a strand of DNA or within the complex of DNA and proteins known as chromatin. By occupying these sites like roadblocks on a highway, methyl groups can alter transcription, a basic step in gene expression where a piece of RNA is made from a DNA template. Increased methylation generally impedes RNA transcription, whereas less methylation enhances transcription. These changes are enduring and persist when cells divide.

In 2015, Yehuda's group became one of the first to find epigenetic changes on stress-related genes of veterans with PTSD. These alterations partially explained why trauma's effects were so persistent, lasting for decades. Specifically, reduced methylation was found in an important region of a gene that encodes the glucocorticoid receptor, likely increasing the sensitivity of these receptors.

This epigenetic modification suggests a potential explanation for how trauma might reset cortisol levels. The body regulates the stress response through a complicated feedback mechanism. A rise in cortisol levels will prompt the body to produce less of the hormone, which may drive up the numbers and responsiveness of glucocorticoid receptors.

Given the epigenetic and other changes occurring with sustained responses to trauma, the feedback loop might become recalibrated. In people who have already endured trauma, their stress systems might be sensitized and their cortisol levels diminished—increasing their adrenaline response to further trauma and leading to PTSD.

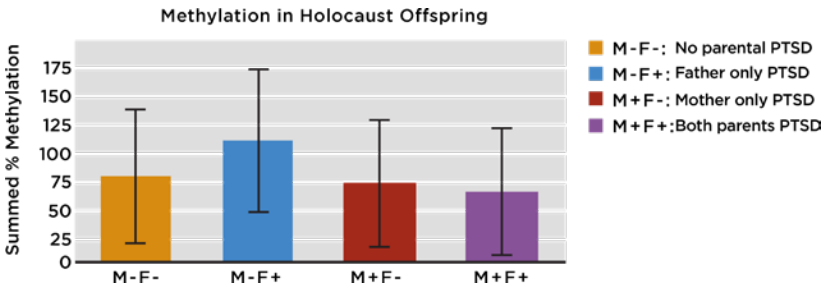
Epigenetic Inheritance through Mothers versus Fathers

Researchers had assumed all along that trauma was behaviorally transmitted. But now, it looked like the uterine environment also played a role.

Early studies of Holocaust offspring had selected only those people with two parents who were Holocaust survivors. The studies were redone to figure out if the sex of the parent mattered—and it did. Those whose mothers had PTSD tended to have lower cortisol levels and showed evidence of more sensitive glucocorticoid receptors. In contrast, those whose fathers, and not mothers, had PTSD showed the opposite effect.

Taking a closer look, researchers again discovered lower methylation within the glucocorticoid receptor gene in Holocaust offspring whose mothers, or both parents, had PTSD. These changes mirrored what had been observed in the maternal survivors themselves.

But in offspring with only paternal PTSD, they observed more methylation—the opposite effect. These findings raised the possibility that PTSD in mothers and fathers might lead to different epigenetic changes on the glucocorticoid receptor in children.



A second series of studies, beginning in 2016, examined methylation within another gene, one that encodes a protein involved in regulating the ability of the glucocorticoid receptor to bind cortisol. The findings showed related methylation patterns in both Holocaust parents and their children.

Researchers were able to extend this work in a larger sample of offspring of Holocaust survivors. In 2020, they reported lower levels of methylation in the adult children whose mothers—and not fathers—were exposed to the Holocaust during childhood. This effect was independent of whether the mother had PTSD. It suggested that trauma might have affected the mothers' eggs decades before her children were conceived, while she was herself a child.

Given the obvious difficulties in studying generations of people, scientists often turn to animal studies to explore epigenetic transmission. In 2014, researchers at Emory University reported an intergenerational epigenetic pathway that ran through sperm. They gave a male mouse a mild electric shock as it smelled a cherry blossom scent, stimulating a fear response to the odor. The response was accompanied by epigenetic changes in its brain and sperm.

Intriguingly, the male offspring of the shocked mice demonstrated a similar fear of cherry blossoms—as well as epigenetic changes in their brain and sperm—without being exposed to the shock. These effects were passed down for two generations, from grandfather to son to grandson.

In Utero Effects of Trauma

Apart from altering the eggs and sperm that encapsulate our genetic inheritance, sometimes decades before conception, trauma also seems to influence the uterine environment. Meticulous studies have followed the offspring of women who were pregnant during the Dutch famine—a 6-month period during World War II when the Nazis blocked the food supply to the Netherlands, causing widespread starvation.

Researchers discovered that the combined effects of extreme stress and nutritional deprivation could cause deficits in metabolism and susceptibility to cardiovascular illness. But the effects depended on the trimester of exposure.

How might the uterine environment leave a trauma trace in the offspring? Work on Holocaust survivors and their adult children provided some clues. The story is again complicated, and it involves an enzyme known as 11-beta-hydroxysteroid dehydrogenase type 2.

Holocaust survivors had lower levels of the enzyme than those who hadn't lived through the Holocaust—and such effects were particularly pronounced in those who were the youngest during World War II. The enzyme is normally concentrated in the liver, kidneys, and brain. Under conditions of food deprivation, however, the body can lower levels of this enzyme to increase metabolic fuel in the interest of promoting survival.

In adults, the enzyme level will return to what it had been originally when there is no more starvation, but in children, the level may remain low. The findings suggested that levels of the enzyme might have been altered during childhood when Holocaust survivors were exposed to long periods of malnourishment, and the change persisted well into old age.

In the children of women who were Holocaust survivors, however, researchers saw quite the opposite: Levels of the enzyme were higher than in Jewish control subjects. The result might seem contradictory, but there is a logic to it. During pregnancy, the enzyme also acts in the placenta, protecting the fetus from exposure to circulating maternal cortisol, which can be toxic to the developing brain. The high levels of this enzyme in the offspring of Holocaust survivors may thus reflect an adaptation, an effort to protect the fetus from the lowered enzyme levels in their mothers.

All of this means that offspring are not always passive recipients of their parents' scars. Just as a parent was able to survive trauma by means of biological adaptations, offspring can sometimes adapt to the biological impact of their parents' trauma.

Permanent Damage or Survival Advantage?

An important question is whether epigenetic alterations in stress-related genes, particularly those reflected in the offspring of traumatized parents, are necessarily markers of vulnerability or whether they may reflect a mechanism through which offspring become better equipped to cope with adversity. This is an area Rachel Yehuda and others are actively exploring.

It is tempting to interpret epigenetic inheritance as a story of how trauma results in permanent damage. Epigenetic influences might nonetheless represent the body's attempts to prepare offspring for challenges similar to those encountered by their parents.

As circumstances change, however, the benefits conferred by such alterations may wane or even result in new vulnerabilities. The survival advantage of this form of intergenerational transmission depends in large part on the environment encountered by the offspring themselves.

Moreover, some of these stress-related and intergenerational changes may be reversible. Several years ago, it was discovered that combat veterans with PTSD who benefited from cognitive-behavioral psychotherapy showed treatment-induced changes in methylation. The finding confirmed that healing is also reflected in epigenetic change. Such findings represent an important frontier in psychiatry and may suggest new avenues for treatment.

ABOUT THIS LECTURE


This content was adapted from the article "Trauma in the Family Tree" by Rachel Yehuda. It appeared in the July 2022 issue of *Scientific American*.

9



The Breakthrough in Messenger RNA Therapies





IN JUST 17 YEARS, MESSENGER RNA (MRNA) therapies went from proof of concept to global salvation. The Pfizer-BioNTech and Moderna vaccines for COVID-19 have been given to hundreds of millions of people, saving countless lives. This lecture looks at how mRNA vaccines and drugs differ from conventional therapies, what diseases researchers hope to target with mRNA therapies, and what challenges lie ahead for worldwide production.

Differences between Messenger RNA and Other Therapies

In 2005, Katalin Karikó and Drew Weissman created a way to make mRNA molecules that would not cause dangerous inflammation when injected into an animal's tissue. In 2017, Weissman and Norbert Pardi demonstrated that modified mRNA, carried into human cells by a fatlike nanoparticle, was protected from being broken down by the body and prompted immune cells to generate antibodies that neutralize an invading virus more effectively than the immune system could on its own.

The COVID vaccines made by Pfizer-BioNTech and Moderna both use this mRNA-liquid-nanoparticle “platform”—known as mRNA-LNP. In large clinical trials, it prevented more than 90% of the people who received the vaccines from becoming ill. The platform outperformed more conventional approaches, in which vaccines are grown in laboratory cell cultures or chicken eggs.

Messenger RNA vaccines instruct cells to create proteins that induce an immune response to an invader such as SARS-CoV-2, training the immune system to attack the actual pathogen if it infects the person in the future. These vaccines are easier to produce in large quantities than conventional protein therapies, which are genetically engineered versions of natural human or pathogen proteins.

They're also easier than monoclonal antibody therapies, which are lab-produced molecules that attack viruses in the same way that human antibodies do. And once a reliable manufacturing facility is built, it can quickly switch to a new mRNA vaccine or drug—unlike protein or monoclonal facilities, which must reengineer production from the ground up for each new therapy.

Messenger RNA Therapies for Other Diseases

Success has inspired researchers, companies, and government labs to pursue mRNA therapies for many infectious diseases, including influenza, cytomegalovirus, herpes simplex virus 2, norovirus, rabies,

malaria, tuberculosis, dengue, Zika, HIV, hepatitis C, and the entire family of coronaviruses. In each case, researchers wanted to determine exactly how mRNA-LNP vaccines induce potent antibody responses.

Work on mRNA vaccines also has expanded to certain cancers, food and environmental allergies, and autoimmune diseases. Positive results against ATTR amyloidosis, a fatal condition that involves the liver, have already been found in a phase 1 clinical trial.

Although the use of protein-based medications for certain illnesses is expanding quickly, large doses are typically required, and production is often difficult and expensive; mRNA delivery of therapeutic proteins could help. The approach has already worked in lab animals for issues including bone repair and asthma, and some treatments have advanced to human clinical trials.

The concentrated COVID work has also helped make mRNA a leader in nucleic acid therapeutics—approaches that can produce nearly any protein made by a specific cell. The technique is starting to be applied, and it could fight diseases in more convenient, less invasive, and less expensive ways.

For example, the FDA has approved gene therapy for sickle cell anemia, and it is working in the US, although it requires marrow to be extracted from a person's bone, treated, and reinserted. Messenger RNA therapy could be delivered to marrow with a straightforward injection into a person's arm. If that works, sickle cell treatment could be greatly expanded in countries where the condition is widespread.

Plenty of hurdles remain, including the creation of a better supply chain for delivering raw mRNA vaccine and materials needed for its production worldwide, as well as improvements that could reduce the dosage a person needs. Yet the ease of mRNA production should enable most countries to make their own medications and develop new therapeutics.

ABOUT THIS LECTURE

This content was adapted from the article “Messenger RNA Therapies Finally Arrived” by Drew Weissman. It appeared in the March 2022 special issue, *How COVID Changed the World*.

10



You Really Can Learn in Your Sleep



EVERYONE KNOWS WE LEARN BETTER WHEN we are well rested. Most people, however, dismiss the notion of sleep learning out of hand. Yet a set of new neuroscientific findings complicates this picture by showing that a critical part of learning occurs during sleep: Recently formed memories resurface during the night, and this playback can help reinforce them, allowing at least a few to be remembered for a lifetime. This lecture looks at the connections between sleep, learning, and memory.



Early Studies of Sleep Learning

In Aldous Huxley's *Brave New World*, a boy memorizes each word of a lecture in English, a language he does not speak. The learning happens as the boy sleeps within earshot of a radio broadcast of the lecture. On awakening, he is able to recite the entire lecture. Based on this discovery, the totalitarian authorities of Huxley's dystopian world adapt the method to shape the unconscious minds of all their citizens.

Around the time that Huxley was writing *Brave New World*, serious explorations into the possibility of meddling with sleep had begun. In 1927, New Yorker Alois B. Saliger invented an Automatic Time-Controlled Suggestion Machine, which he marketed as the PsychoPhone, to allow a recorded message to be replayed during the night. The setup seemed to evoke Huxley's imagined technology except that the user, rather than the state, could select the message to be played.

The field took a severe hit in 1956, when two scientists at RAND Corporation used electroencephalography (EEG) to record brain activity while 96 questions and answers were read to sleeping study participants. The next day, correct answers were recalled only for information presented when sleepers showed signs of awakening.

These results led to a shift in the field that persisted for 50 years, as researchers began to lose faith in sleep learning as a viable phenomenon: Participants in these experiments appeared to learn only if they were not really sleeping when information was presented to them.

Most scientists during this time tended to avoid the topic of sleep learning, although a few researchers did plug away at asking whether sleep assists in remembering new information. One typical study protocol probed whether overnight sleep deprivation affected recall the day after learning something new. Another asked whether remembering was better after a nap than after an equal period of time spent awake.

Various confounding factors can interfere with such studies. For example, the stress of sleep deprivation can harm cognitive functions, which then decreases memory recall. Eventually, cognitive neuroscientists began to tackle these challenges by bringing together evidence from multiple research methods.



Slow-Wave Sleep and Memory

A substantive foundation of evidence gradually confirmed that sleep is a means of reviving memories acquired during the day and helping the brain store the memories. This reopened the relation between sleep and memory as a legitimate area of scientific study.

Many researchers who took up the challenge focused on rapid eye movement (REM) sleep, the period when dreams are the most frequent and vivid. The guiding assumption held that the brain's nighttime processing of memories would be tied to dreaming, but clear-cut data did not materialize.

Instead of REM, some investigators focused their attention on slow-wave sleep (SWS), a period of deep slumber without rapid eye movements. In 2007, Björn Rasch, then at the University of Lübeck in Germany, and his colleagues prepared people for a sleep experiment by requiring them to learn the locations of a set of objects while simultaneously smelling the odor of a rose.

Later, in their beds in the laboratory, sleeping study participants again encountered the same odor as electrical recordings confirmed one sleep stage or another. The odor activated the hippocampus, a brain area critical for learning to navigate one's surroundings and for storing the new knowledge. On awakening, participants recalled locations more accurately—but only following cueing from odors they encountered during the course of SWS.

Targeted Memory Reactivation

In 2009, Ken A. Paller and Delphine Oudiette at Northwestern University extended this methodology by using sounds instead of odors. They found that sounds played during SWS could improve recall for individual objects.

In their procedure—termed targeted memory reactivation—they first taught people the locations of 50 objects. They might learn to place a cat at one designated spot on a computer screen and a teakettle at another. At the same time, they would hear a corresponding sound (a meow for the cat, a whistle for the kettle, and so on).

After this learning phase, participants took a nap in a comfortable place in the lab. EEG recordings provided intriguing data on the synchronized activity of networks of neurons in the brain's outer layer, the cerebral cortex, that are relevant for memory reactivation.

When the scientists detected SWS, they softly played the sounds associated with a subset of the objects from the learning phase. On awakening, people remembered locations cued during sleep better than locations that had not been cued during sleep. Whether sounds or odors served as cues in these experiments, they apparently triggered the reactivation of spatial memories and reduced forgetting.

Besides boosting spatial memory, these procedures have also helped improve recall in other settings. Targeted memory reactivation can assist in mastery of playing a keyboard melody and learning new vocabulary or grammatical rules.

Paller and Oudiette, together with Northwestern University neurologist Marc Slutzky, are testing a rehabilitation procedure for recovering arm-movement abilities after stroke. Cue sounds are incorporated as part of the therapy and are replayed during sleep to try to accelerate relearning of lost movements.

Creating Memories during Sleep

The ability to reinforce memories raises the question of whether new information can be loaded into a person's brain after falling asleep, a technique that calls forth the ethical specter of mind control invoked by *Brave New World*. Is it going too far, though, to imagine that memories can be created surreptitiously?

Although the orthodox response to such conjectures has for many years been an unqualified no, studies at the Weizmann Institute of Science in Israel demonstrated the creation of relatively simple memories using odors.

In one experiment, the researchers succeeded in diminishing the desire for tobacco in smokers who were keen to quit. When asleep, study participants were exposed to two odors, cigarette smoke and rotten fish. During the next week, those who had smelled the mix of both odors smoked 30% less, having apparently been conditioned to associate smoking with the aversive fish odor.

In a collaboration by researchers in France, Germany, the Netherlands, and the US, researchers used a variant of the targeted memory reactivation method to encourage lucid dreaming—a state in which people realize they are dreaming while remaining in the midst of the dream.

They then showed that people could understand softly spoken questions and produce correct answers by signaling with their eyes, their breathing, or subtle twitches in their facial muscles. Sometimes, people in these experiments woke up able to recollect parts of their dreamtime Q and A. These rare occurrences document full-blown learning experienced during sleep.

Looking Ahead

Sleep is needed not just to help people stay alert and rejuvenated but also to reinforce memories acquired while they were awake. We still need to learn much more about offline memory processing. Further work must ascertain how sleep helps learning. It is also essential to find out more about the perils of poor or inadequate sleep.

All this knowledge might help in creating programs of sleep learning to preserve memories, to speed the acquisition of new knowledge, or even to change bad habits, such as smoking. Looking still further ahead, scientists might also explore whether we can gain control over our dreams, which could lead to the prospect of nightmare therapies, sleep-based problem-solving, and perhaps even recreational dream travel.

In a culture that already offers wrist-based sleep trackers, we can contemplate new ways to convert sleep into a productive endeavor—for some, a chilling prospect, and for others, another opportunity for self-improvement.

ABOUT THIS LECTURE

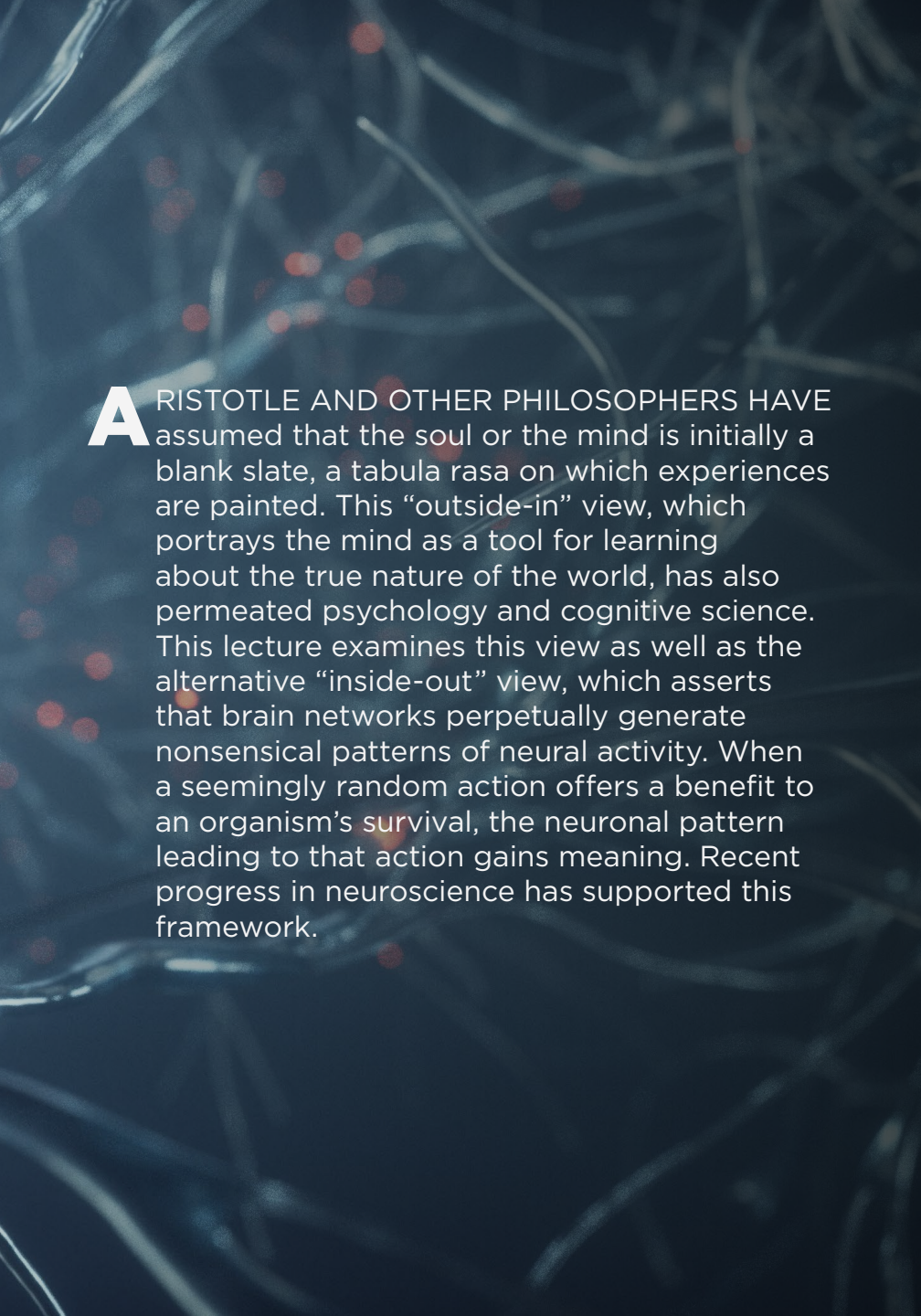
This content was adapted from the article “Sleep Learning Gets Real” by Ken A. Paller and Delphine Oudiette. It appeared in the February 2022 collector’s edition, *Secrets of the Mind*.



11



How Your Brain Constructs the World



ARISTOTLE AND OTHER PHILOSOPHERS HAVE assumed that the soul or the mind is initially a blank slate, a tabula rasa on which experiences are painted. This “outside-in” view, which portrays the mind as a tool for learning about the true nature of the world, has also permeated psychology and cognitive science. This lecture examines this view as well as the alternative “inside-out” view, which asserts that brain networks perpetually generate nonsensical patterns of neural activity. When a seemingly random action offers a benefit to an organism’s survival, the neuronal pattern leading to that action gains meaning. Recent progress in neuroscience has supported this framework.

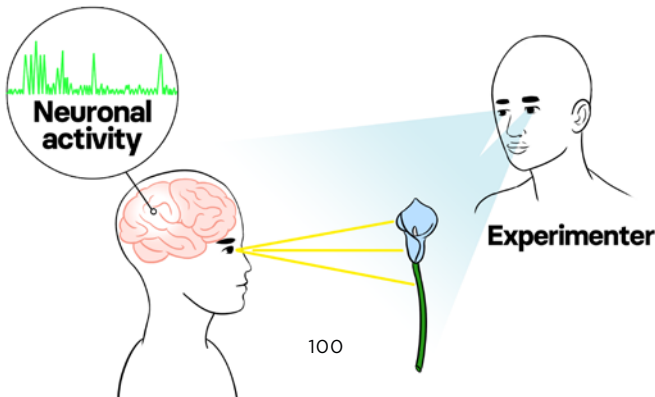
The Outside-In Framework

The dominance of the blank slate framework is illustrated by the Nobel Prize–winning experiments by David Hubel and Torsten Wiesel. In their signature experiments, they recorded neural activity in animals while showing them images of various shapes. Moving lines, edges, light or dark areas, and other physical qualities caused different sets of neurons to fire.

The assumption was that neuronal computation starts with simple patterns that are synthesized into more complex ones. These features are then bound together somewhere in the brain to represent an object. No active participation is needed. The brain automatically performs this exercise.

This outside-in framework presumes that the brain’s fundamental function is to perceive “signals” from the world and correctly interpret them. But if this assumption is true, an additional operation is needed to respond to these signals. Wedged between perceptual inputs and outputs is a hypothetical central processor—which takes in sensory representations from the environment and makes decisions about what to do with them to perform the correct action.

So, what exactly is the central processor in this outside-in paradigm? This speculative entity goes by various names—free will, homunculus, decision-maker, executive function, intervening variables, or simply just a “black box.” It all depends on the experimenter’s philosophical inclination and whether the mental operation in question is applied to the human brain, brains of other animals, or computer models. Yet all these concepts refer to the same thing.



Neuroscientists are trying to find where the putative central processor resides in the brain and describe the neuronal mechanisms of decision-making. The physiology of decision-making has become one of the most popular fields in contemporary neuroscience. Higher-order brain regions, such as the prefrontal cortex, have been postulated as the place where “all things come together” and “all outputs are initiated.”

Assigning Meaning to Neuronal Changes

When we look more closely, however, the outside-in framework does not hold together. This approach cannot explain how photons falling on the retina are transformed into a recollection of a summer outing. The outside-in framework requires the artificial insertion of a human experimenter who observes this event.

The experimenter-in-the-middle is needed because even if neurons change their firing patterns when receptors on sensory organs are stimulated—by light or sound, for instance—these changes do not intrinsically “represent” anything that can be absorbed and integrated by the brain. There is no interpreter in the brain to assign meaning to the changes.

Without a magical homunculus watching the activities of all the neurons in the brain with the omniscience of the experimenter, the neurons that take this all in are unaware of the events that caused these changes in their firing patterns.

Because neurons have no direct access to the outside world, they need a way to compare or “ground” their firing patterns to something else. The term *grounding* refers to the ability of the brain’s circuits to assign meaning to changes in neuronal firing patterns that result from sensory inputs.

They accomplish this task by relating this activity to something else. The “dah-dah-dit” Morse code pattern becomes meaningful only when it has previously been linked to the letter “G.” In the brain, the only available source of a second opinion appears when we initiate some action.

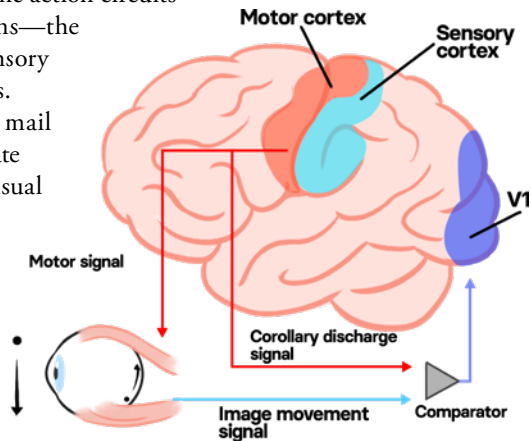
For example, the distances between two trees and two mountain peaks may appear identical, but by moving around and shifting our perspective, we learn the difference.

Corollary Discharges

The outside-in framework follows a chain of events from perception to decision to action. In this model, neurons in dedicated sensory areas are “driven” by environmental signals and thus cannot relate their activity to something else. But the brain is not a serial processing unit; it does not proceed one by one through each of these steps. Instead, any action a person takes involves the brain’s motor areas informing the rest of the cerebral cortex about the action initiated—a message known as a corollary discharge.

Neuronal circuits that initiate an action dedicate themselves to two tasks. The first is to send a command to the muscles that control the eyes and other bodily sensors (the fingers and tongue, among others). These circuits orient bodily sensors in the optimal direction and enhance the brain’s ability to identify the nature and location of signals from the senses.

The second task of these same action circuits involves sending notifications—the corollary discharges—to sensory and higher-order brain areas. Think of them as registered mail receipts. Neurons that initiate eye movement also notify visual sensory areas of the cortex about what is happening and disambiguate whether, say, a flower is moving in the wind or being handled by the person observing it.



This corollary message provides the second opinion sensory circuits need for grounding—a confirmation that “my own action is the agent of change.” Similar corollary messages are sent to the rest of the brain when a person takes actions to investigate the flower and its relationship to that person and other objects.

Without such exploration, visual stimuli from the flower alone would never become signals that furnish a meaningful description of the flower’s size and shape. Perception then can be defined as what we do—not what we passively take in through our senses.

Learning through Matching

The contrast between outside-in and inside-out approaches becomes most striking when used to explain learning. A tacit assumption of the blank slate model is that the complexity of the brain grows with the amount of experience. As we learn, the interactions of brain circuits should become increasingly more elaborate.

In the inside-out framework, however, experience is not the main source of the brain’s complexity. Instead, the brain organizes itself into a vast repertoire of patterns of firing known as neuronal trajectories. This self-organized brain model can be likened to a dictionary filled initially with nonsensical words. Learning takes place through a process of matching the preexisting neuronal trajectories to events in the world.

To understand the matching process, we need to examine the advantages and constraints brain dynamics impose on experience. In the basic version, models of blank slate neuronal networks assume a collection of largely similar, randomly connected neurons. The presumption is that brain circuits are highly plastic and that any arbitrary input can alter them.

We can see a problem with this approach by considering an example from the field of artificial intelligence. Artificial neural networks built to “write” inputs onto a neural circuit often fail because each new input inevitably modifies the circuit’s connections and dynamics. The circuit is said to exhibit plasticity.

But there is a pitfall. While constantly adjusting the connections in its networks when learning, the AI system, at an unpredictable point, can erase all stored memories—a bug known as catastrophic interference, an event a real brain never experiences.

The inside-out model, in contrast, suggests that self-organized brain networks should resist such perturbations. Yet they should also exhibit plasticity selectively when needed. The way the brain strikes this balance relates to vast differences in the connection strength of different groups of neurons.

The Roles of Different Neuronal Connections

Connections among neurons exist on a continuum. Most neurons are only weakly connected to others, but a smaller subset retains robust links. The strongly connected minority is always on the alert. It fires rapidly, shares information readily within its own group, and stubbornly resists any modifications to the neurons' circuitry.

These elite subnetworks, sometimes described as a “rich club,” remain well informed about neuronal events throughout the brain. The rich club makes up roughly 20% of the overall population of neurons, but it is in charge of nearly half of the brain's activity.

In contrast to the rich club, most of the brain's neurons—the neural “poor club”—tend to fire slowly and are weakly connected to other neurons. But they are also highly plastic and able to physically alter the connection points between neurons, known as synapses.

Both rich and poor clubs are key for maintaining brain dynamics. Members of the ever-ready rich club fire similarly in response to diverse experiences. They offer quick, good-enough solutions under most conditions. We make good guesses about the unknown not because we remember it but because our brains always make a guess about a new, unfamiliar event.

Nothing is completely novel to the brain because it always relates the new to the old. It generalizes. Even an inexperienced brain has a vast reservoir of neuronal trajectories at the ready, offering opportunities to match events in the world to preexisting brain patterns without substantial changes to connections. A brain that remakes itself constantly would be unable to adapt quickly to fast-changing events in the outside world.

But there also is a critical role for the plastic, slow-firing-rate neurons. These neurons come into play when something of importance is detected and needs to be recorded for future reference. They then go on to mobilize their vast reserve to capture subtle differences between one thing and another by changing the strength of some connections to other neurons.

Creating Our Internal and External Worlds

Neurons devote most of their activity to sustaining the brain's internal states rather than being controlled by sensations. During the course of natural selection, organisms adapt to the ecological niches in which they live and learn to predict the likely outcomes of their actions in those niches. As brain complexity increases, more intricate connections and neuronal computations insert themselves between motor outputs and sensory inputs.

This investment enables the prediction of planned actions in more complex and changing environments and at lengthy timescales far in the future. More sophisticated brains also organize themselves to allow computations to continue when sensory inputs vanish temporarily and an animal's actions come to a halt.

When you close your eyes, you still know where you are because a great deal of what defines "seeing" is rooted in brain activity. This disengaged mode of neuronal activity provides access to a virtual world of experience.

John O'Keefe of University College London won a Nobel Prize for discovering that firing of an animal's hippocampal neurons during navigation correlates with that animal's spatial location. For that reason, these neurons are known as place cells.

When a rat walks through a maze, distinct assemblies of place cells become active in a sequence corresponding to where it is on its journey. From that observation, one can tentatively conclude that sensory inputs from the environment control the firing of neurons, in line with the outside-in model.

Yet other experiments, including in humans, show that these same networks are used for our internal worlds that keep track of personal memories, engage in planning, and imagine future actions. If cognition is approached from an inside-out perspective, it becomes clear that navigation through either a physical space or an imaginary landscape is processed by identical neural mechanisms.

Episodic memories are more than recollections of past events. They also let us look ahead to plan for the future. They function as a kind of “search engine” that allows us to probe both past and future. This realization also suggests the nomenclature should be broadened. These experiments show that progressions of place cell activity are internally generated as preconfigured sequences selected for each maze corridor. Same mechanism, multiple designations—so they can be termed place cells, memory cells, or planning cells, depending on the circumstance.

Disengaged Brain Activity

Further support comes from “offline” brain activity when an animal is either doing nothing, consuming a reward, or just sleeping. As a rat rests in the home cage after a maze exploration, its hippocampus generates brief, self-organized neuronal trajectories.

These sharp wave ripples, as they are known, repeat the neuronal sequences that occurred during maze traversals. Sharp wave-ripple sequences help to form our long-term memories and are essential to normal brain functioning.

Clever experiments performed in human subjects and in animals over the past decade show that the time-compressed ripple events constitute an internalized trial-and-error process that subconsciously creates real or fictional alternatives for making decisions about an optimal strategy, constructing novel inferences, and planning ahead for future actions without having to immediately test them.

In this sense, our thoughts and plans are deferred actions, and disengaged brain activity is an active, essential brain operation. In contrast, the outside-in theory does not make any attempt to assign a role to the disengaged brain when it is at rest or even in the midst of sleep.

In real brains, neural processes that operate through disengagement from the senses go hand in hand with mechanisms that promote interactions with the surrounding world. All brains, simple or complex, use the same basic principles. Disengaged neural activity, calibrated simultaneously by outside experience, is the essence of cognition.

ABOUT THIS LECTURE

This content was adapted from the article "Constructing the World from Inside Out" by György Buzsáki. It appeared in the June 2022 issue of *Scientific American*.

12



Why Black Holes Turn Inside Out

THEORETICAL PHYSICS ENTERED CRISIS mode in 1974, when Stephen Hawking argued that black holes destroy information. Hawking showed that a black hole can evaporate, gradually transforming itself and anything it consumes into a featureless cloud of radiation. During the process, information about what fell into the black hole is apparently lost, and irretrievably losing information violates a sacred principle of physics. This problem remained open for almost 50 years, but the pieces started falling into place in 2019. The resolution is based on a new understanding of space-time and how it can be rewired through quantum entanglement, which leads to the idea that part of the inside of a black hole, the so-called island, is secretly on the outside. This lecture looks at how scientists arrived at these new ideas.

The Event Horizon

Black holes are formed when enough matter is confined within a small enough region that space-time collapses in on itself in a violent feedback loop of squeezing and stretching that fuels more squeezing and stretching.

These tidal forces cause the abrupt end of an entire region of space-time at the so-called black hole singularity—the place where time stops and space ceases to make sense. There is a fine line within the collapsing region that divides the area where escape is possible from the point of no return. This line is called the event horizon. It is the outermost point from which light barely avoids falling into the singularity.

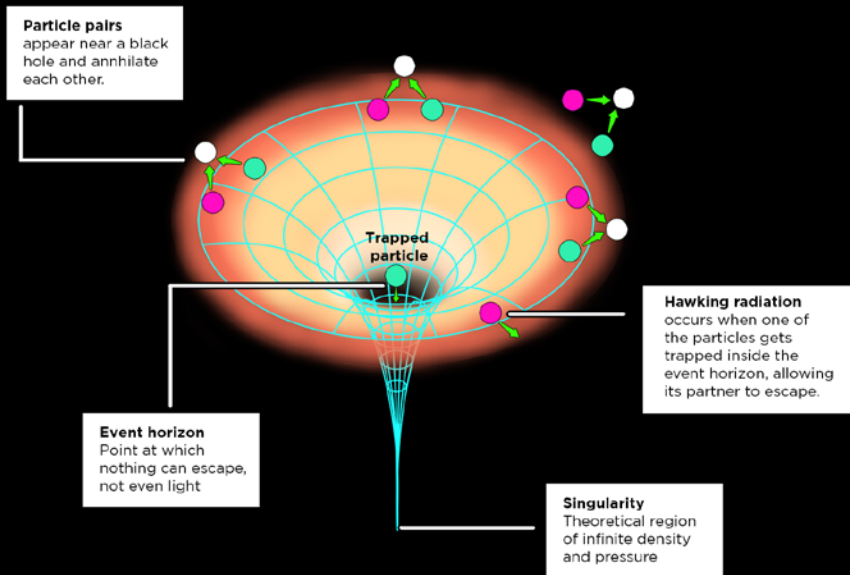
Unless a thing travels faster than light—which is a physical impossibility—it cannot escape from behind the event horizon; it is irretrievably stuck inside the black hole. The one-way nature of this boundary is not immediately problematic. In fact, Einstein's general theory of relativity predicts it. The danger starts when this notion interacts with the wild world of quantum mechanics.

Hawking Radiation

Quantum theory redeems black holes from being the greedy monsters they are made out to be. Every calorie of energy they consume they eventually give back in the form of Hawking radiation—energy squeezed out of the vacuum near the event horizon.

In quantum theory, a sea of particles—photons, electrons, gravitons, and more—make up empty space. These particles come in carefully arranged pairs, acting hand in hand as the glue that holds space-time together.

Particle pairs that straddle the event horizon of a black hole, however, become forever separated from each other. The newly divorced particles peel away from the horizon in opposite directions, with one member crashing into the singularity and the other escaping the black hole's gravitational pull in the form of Hawking radiation.



This process is draining the black hole, causing it to get lighter and smaller as it emits energy in the form of the outgoing particles. Because of the law that energy must be conserved, the particles trapped inside must then carry negative energy to account for the decrease in the total energy of the black hole.

From the outside, the black hole appears to be slowly burning away. When you burn a book, the words on its pages imprint themselves on the pattern of the emanating light and the remaining ashes. This information is thus preserved, at least in principle.

If an evaporating black hole were a normal system like the burning book, then the information about what falls into it would be encoded into the emerging Hawking radiation. Unfortunately, this is complicated by the quantum-mechanical relation among the particles across the horizon.

The issue begins with the end of the pairing of the two particles straddling the event horizon. Despite being separated, they maintain a quantum union that transcends space and time—they are connected by entanglement. Quantum entanglement is perhaps one of the weirdest aspects of our universe and arguably one of its most essential.

Entanglement

Here's a simple example of entanglement. Consider two coins in what's called superposition—the quantum phenomenon of being in multiple states until a measurement is made. In this case, it's a superposition of both coins being either heads or tails. The coins aren't facing heads and tails at the same time—that's physically impossible—but the superposition means that the chance of observing the pair of coins in either orientation, both heads or both tails, is a probability of one half.

There is no chance of ever finding the coins in opposite orientations. The two coins are entangled; the measurement result of one predicts the result of the other with complete certainty. Either coin by itself is completely random, devoid of information, but the randomness of the pair is perfectly correlated.

The scientists who described this phenomenon were troubled by how the two coins appeared to influence each other without having to come into physical contact. The coins could be in separate galaxies while still maintaining the same amount of entanglement between them. Einstein doubted the reality of what he called the “spooky action at a distance” linking the results of the two separate random measurements.

The irony is that Einstein himself is in a superposition of being both wrong and right. He was right to recognize the importance of entanglement in distinguishing quantum mechanics from classical physics. What he got wrong can be summed up with the truism “correlation does not imply causation.”

Although the fates of the particles are inextricably correlated, the measurement outcome of one does not cause the outcome of the other. It turns out that quantum mechanics simply allows for a new, higher degree of correlation than we are used to.

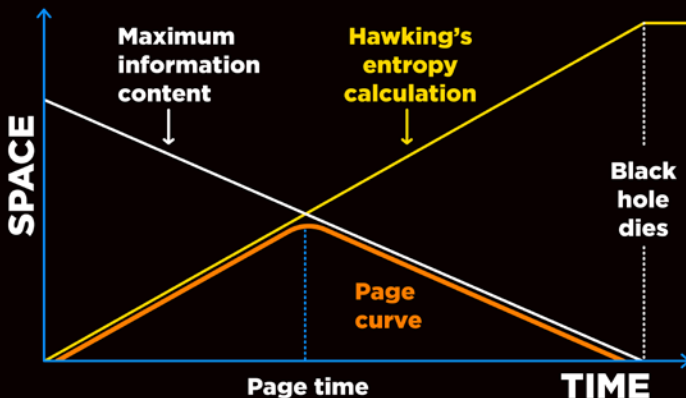
Entanglement Entropy and the Page Curve

Because Hawking radiation is composed of one half of a collection of entangled pairs, it emerges from the black hole in a completely random state—if the particles were coins, they would be observed to be heads or tails with equal probability. Hence, we cannot infer anything useful about the contents of the black hole from the random measurements of the radiation.

We can measure the lack of information—or the randomness—in the Hawking radiation by thinking about the amount of entanglement between the radiation and the black hole. This is because one member of an entangled pair is always random, and the outside members are all that remains by the end of the evaporation.

The calculation of randomness goes by many names, including entanglement entropy, and it grows with every emerging Hawking particle, plateauing at a large value once the black hole has completely disappeared.

This pattern differs from what happens when information is preserved, as in the example of a burning book. In such a case, the entropy may rise initially, but it has to peak and then fall to zero by the end of the process. This rising and falling pattern of entropy, known as the Page curve, applies to all normal quantum-mechanical systems. The time at which the entropy peaks and then starts to decrease is the Page time.



The Information and Firewall Paradoxes

The destruction of information inside black holes spells disaster for physics because the laws of quantum mechanics stipulate that information cannot be obliterated. This is the famous information paradox—the fact that a sprinkling of quantum mechanics onto the description of black holes leads to a seemingly insurmountable inconsistency.

Physicists knew we needed a more complete understanding of quantum-gravitational physics to generate the Page curve for the Hawking radiation. Unsurprisingly, this task proved difficult.

Part of the challenge was that no minor tweaking of the evaporation process was sufficient to generate the Page curve and send the entropy back down to zero. What we needed was a drastic reimagining of the structure of a black hole.

In a paper published in 2013 by Ahmed Almheiri, Donald Marolf, Joseph Polchinski, and Jamie Sully (known collectively as AMPS), they tried out several ways to modify the picture of evaporating black holes using a series of gedankenexperiments—the German term for the kind of thought experiments Einstein popularized. They concluded that to save the sanctity of information, one of two things had to give: Either physics must be nonlocal—allowing for information to instantaneously disappear from the interior and appear outside the event horizon—or a new process must kick in at the Page time.

To preclude the increase of entropy, this process would have to break the entanglement between the particle pairs across the event horizon. The former option—making physics nonlocal—was too radical, so the researchers decided to go with the latter.

This modification helps to preserve information, but it poses another paradox. Recall that the entanglement across the horizon was a result of having empty space there—because the vacuum of empty space is maintained by a sea of entangled pairs of particles. The entanglement is key; breaking it comes at the cost of creating a wall of extremely high-energy particles, which the AMPS group named the firewall.

Having such a firewall at the horizon would forbid anything from entering the black hole. Instead, infalling matter would be vaporized on contact. The black hole at the Page time would suddenly lose its interior, and space-time would come to an end, not at the singularity deep inside the black hole but right there at the event horizon. This conclusion is known as the firewall paradox, a catch-22 that meant any solution to the information paradox must come at the cost of destroying what we know about black holes. It was a quagmire.

Wormhole Bridges

Eventually, the AMPS group realized that both the information paradox and the newer firewall paradox arose because attempts to meld quantum mechanics and black hole physics were too timid. It wasn't enough to apply quantum mechanics to only the matter present in black holes—they had to devise a quantum treatment of the black hole space-time as well.

To consider the quantum nature of space-time, the AMPS group relied on a technique designed by Richard Feynman, called the path integral of quantum mechanics. The idea is based on the weird truth that, according to quantum theory, particles don't simply travel along a single path from point A to point B—they travel along a superposition of all the different paths connecting the two points.

Similarly, a quantum space-time can be in a superposition of different complicated shapes evolving in different ways. For instance, if we start and end with two regular black holes, the quantum space-time within them has a nonzero probability of creating a short-lived wormhole that temporarily bridges their interiors.

Usually, the probability of this happening is vanishingly slim. When we carry out the path integral in the presence of the Hawking radiation of multiple black holes, however, the large entanglement between the Hawking radiation and the black hole interiors amplifies the likelihood of such wormholes. This realization came to physicist Ahmed Almheiri through work he did in 2019 with several collaborators, and it was also found by an independent team.

Swapping Black Hole Interiors

Why does it matter if some black holes are connected by wormholes? It turns out that they modify the answer of how much entanglement entropy there is between the black hole and its Hawking radiation. The key is to measure this entanglement entropy in the presence of multiple copies of the system. This is known as the replica trick.

The relevant physical effect of these temporary wormholes is to swap out the interiors among the different black holes. This happens literally: What was in one black hole gets shoved into one of the other copies far away, and the original black hole assumes a new space-time interior from a different one. The swapped region of the black hole interior is called the island, and it encompasses almost the entire interior up to the event horizon.

The swapping is exactly what was needed! Focusing on one of the black holes and its Hawking radiation, the swapped-out island takes with it all the partner particles that are entangled with the outgoing Hawking radiation, and hence, technically, there is no entanglement between the black hole and its radiation.

Including this potential effect of wormholes produces a new formula for the entanglement entropy of the radiation. Instead of Hawking's original calculation, which simply counts the number of Hawking particles outside a black hole, the new formula curiously treats the island as if it were outside and a part of the exterior Hawking radiation. Therefore, the entanglement between the island and the exterior should not be counted toward the entropy.

Instead, the entropy that it predicts comes almost entirely from the probability of the swap actually occurring, which is equal to the area of the boundary of the island—roughly the area of the event horizon—divided by Newton's gravitational constant. As the black hole shrinks, this contribution to the entropy decreases. This is called the island formula for the entanglement entropy of the Hawking radiation.

Solving Two Paradoxes with the Island Formula

The final step in computing the entropy is to take the minimum between the island formula and Hawking's original calculation. This gives us the Page curve that we've been after. Initially, we calculate the entanglement entropy of the radiation with Hawking's original formula because the answer starts off smaller than the area of the event horizon of the black hole.

But as the black hole evaporates, the area shrinks, and the new formula takes the baton as the true representative of the radiation's entanglement entropy. What is remarkable about this result is that it solves two paradoxes with one formula. It appears to address the firewall paradox by supporting the option of nonlocality that the AMPS group originally dismissed.

Instead of breaking the entanglement at the horizon, we are instructed to treat the inside—the island—as part of the outside. The island itself becomes nonlocally mapped to the outside. And the formula solves the information paradox by revealing how black holes produce the Page curve and preserve information. Information can escape a black hole not by surmounting the insurmountable event horizon but by simply falling deeper into the island. Scientists are excited about this breakthrough and have only begun to explore the implications of space-time wormholes and the island formula.

ABOUT THIS LECTURE

This content was adapted from the article "Black Holes, Wormholes and Entanglement" by Ahmed Almheiri. It appeared in the September 2022 issue of *Scientific American*.

